



Prebiotic Evolution and Astrobiology

生命的来历

前生物进化与太空生物学

王子晖　安东尼奥·拉兹卡诺　编著
陈　珊　译





(Q-2580.0101)



Prebiotic Evolution and Astrobiology

生命的来历

前生物进化与太空生物学

作为国内第一本前生物进化与天体生物学领域的教材类书籍，本书系统地探讨了生命的本质和特征，详细介绍了太阳系中前生物探索的范围、方式和最新进展，并深入地从分子和细胞水平分析了前生物进化的本质和方式，同时讨论了现存的问题和有关全生物共同祖先的假说。本书对前生物进化与天体生物学的介绍角度多样、内容详尽、结构严谨。虽然是专业书籍，涉及多方面的相关知识，但是论述深入浅出、便于理解，适合各相关领域的学生阅读，也可帮助有兴趣的普通读者了解相关理论、研究和进展。

本书目录

www.sciencep.com



ISBN 978-7-03-029045-8

科学出版社 科爱传播
http://www.kbooks.cn
editor@kbooks.cn

销售分类建议：生物学、天文学

本版本只限于在中华人民共和国境内销售

定价：68.00 元

Prebiotic Evolution and Astrobiology

生命的来历：前生物进化与太空生物学

王子晖　安东尼奥·拉兹卡诺　编著

陈　珊　译

科　学　出　版　社

北　京

图字:01-2010-3045号

内 容 简 介

作为国内第一本前生物进化与天体生物学方面的教材类书籍，本书系统地探讨了生命的本质和特征，详细介绍了太阳系中前生物探索的范围、方式和最新进展，并深入地从分子和细胞水平分析了前生物进化的本质和方式，同时讨论了现存的问题和有关全生物共同祖先的假说。

本书对前生物进化与天体生物学的介绍角度多样、内容详尽、结构严谨。虽然是专业书籍，涉及多方面的相关知识，但是论述深入浅出、便于理解，在适合各相关领域学生学习的同时，也可帮助有兴趣的普通读者了解相关理论、研究和进展。

This is a translation version of Prebiotic Evolution and Astrobiology
J. Tze-Fei Wong and Antonio Lazcano
The original English language work has been published by Landes Bioscience.
Austin, Texas, U. S. A.

ISBN:978-1-58706-330-5

图书在版编目(CIP)数据

生命的来历:前生物进化与太空生物学/王子晖等编著;陈珊译. —北京:科学出版社,2011

ISBN 978-7-03-029045-8

Ⅰ.①生… Ⅱ.①王… ②陈… Ⅲ.①生物-进化-研究 ②天体生物学-研究 Ⅳ.①Q11 ②Q693

中国版本图书馆 CIP 数据核字 (2010) 第 184351 号

责任编辑:孙红梅 李小汀 刘 晶/责任校对:朱光兰
责任印制:钱玉芬/封面设计:耕者设计工作室

科学出版社 出版
北京东黄城根北街16号
邮政编码 100717
http://www.sciencep.com
中国科学院印刷厂 印刷
科学出版社发行 各地新华书店经销
*
2011年1月第 一 版 开本:B5(720×1000)
2011年1月第一次印刷 印张:16 插页:6
印数:1—2 000 字数:337 000

定价:68.00元

(如有印装质量问题，我社负责调换)

编者简介

王子晖（J. Tze-Fei Wong）从事基因密码研究工作超过 30 年，包括遗传密码的诞生、遗传密码与氨基酸合成途径的共进化、遗传密码的突变可以令其解读非自然氨基酸，以及将遗传密码普遍性的起源追溯至近产甲烷生物的全生命的根。早年在多伦多大学及近年在香港科技大学，他发展出用 tRNA 及氨酰 tRNA 合成酶作为系统树的探针应用方法。他是《酶机制动力学》(*Kinetics of Enzyme Mechanisms*）一书的作者，此书主要探讨了多底物的酶反应。

安东尼奥·拉兹卡诺（Antonio Lazcano）是拉丁美洲最杰出的进化学者，在生命起源及早期生命进化方面是世界领先的权威。他出生在墨西哥，从事关于生命起源方面的工作超过 30 年，著作包括《奇迹般的细菌》(*La Bacteria Prodigiosa*)、《生命的火花》(*La Chispa de la Vida*）及《生命起源》(*El Origen da la Vida*）。他是高等生命起源研究会议中第一位拉丁美洲裔主席，亦担任美国宇航局等多个国际组织的咨询委员。他曾获米兰大学授予的荣誉博士学位（2008)、生物学研究金章（Puebla，1990)、基多圣弗朗西斯科大学创办人奖章（2005）及意大利天文生物学会的弗朗西斯科·雷迪奖（2008)。

中 文 序

也许是人类出于自己的孤独感，而在漫漫宇宙中尽力呼唤远邻的生命同类；也许是人类出于固有的好奇心，而在渺渺太空中极力寻找地球之外也许存在的生命。人类对天体生物的遐想和探索已持续了几千年。

20 世纪被誉为“物理世纪”，其最大的辉煌是给人类插上了翅膀，高飞而远望；20 世纪更催生了“生物世纪”，其最大的影响也许是核酸作为地球上所有生命的遗传物质的证明、DNA 双螺旋结构的阐明，以及伴随“人类基因组计划”而来的、基于基因组序列的“基因组时代”。我们对生命本质的理解日新月异，对生命起源和进化的认识突飞猛进。太阳系以外的行星的发现、火星上水的证实，一次又一次唤起了人类对太空生物学研究的激情。人类对太空的探索中，地球以外可能的生命的存在已成为最前沿的科学课题。

遗传密码是人类第一个系统破译的宇宙语言。作为一名生物系的学生，我在大学时代便对遗传密码情有独钟。我曾经一度沉醉于那 4 种核苷酸和 20 种氨基酸之间的化学结构关系，唯一的收获是迄今还能倒背如流的遗传密码表。我对本书作者王子晖教授的敬佩就是这样开始的。他在 30 多年前提出了遗传密码同进论，解释了遗传密码和氨基酸的共同进化，并首先试图改写 30 多亿年来亘古不变的遗传密码，被称为“合成生物学”的开拓者之一，被英国《新科学家》杂志誉为书写了“地球生命史的下半集”。这也是我斗胆以晚辈之身份为前辈之巨著作序的唯一原因。

安东尼奥·拉兹卡诺教授是生命起源领域的先驱之一。他曾与最先实现前生命式氨基酸化学合成的斯坦利·米勒教授共同提出了生命异养起源说。这两位大师，更邀请了这一领域的其他出色的研究者们，在本书中撰文介绍了前生物进化和天体生物学的最新研究成果。我特别高兴见到本书的中译本，使我国的年轻学生、学者也能同步受益。

我更加期待有这么一天，有这么一个人，会说这么一句话：我在前生物进化和天体生物学这样一个领域的这一贡献，首先要感谢这么一本书。

此为序，为敬，为荐。

杨焕明

庚寅端午

前　言

随着基因组研究及太空探索的快速发展，前生物进化及天体生物学领域得到前所未有的发展，因此需要有关的教科书。由于内容涵盖多方面，因此需要多位专家合力出版，使本书适合不同领域的学生。这本书源于2006年意大利艾托里马约拉纳基金会与科学文化中心在西西里艾里斯举办的一次关于生命起源的基本问题的会议。在那里，Pier Luigi Luisi对这个主题制订了一个平衡的计划，这本书基本上源于这个计划。

全书共分15章。第1章论述了细胞生物学和分子生物学基础，讨论了生命的本质和生命的确定时刻这个古老命题。第2章以最小细胞的特征带出前生物进化的目标。第3～5章介绍了目前太空生物学探索的范围，主要集中在木卫二、土卫六、土卫二、火星及彗星。由于2008年历史性地在火星上发现水，使这个在21世纪最遥远的科学前沿领域变得更特殊。第6～7章集中讨论陨星的化学及左旋性手性氨基酸在陨石中的出现，揭示了令人意外的太空化学和生物化学共享的特征。由于早期地球是前生物进化的舞台，它的条件对于生命的出现是非常重要的，这会在第8章中详述。第9～10章讨论了前生物合成中利用大大小小的生物分子直到RNA的出现。第11章探索RNA世界的内部系统。第12章描述功能性原始细胞的原理及前细胞进化的不同模式。tRNA及蛋白质的断裂基因之谜在第13章进行分析。基因密码的起源、进化及突变在第14章讨论。第15章考查生命之根，即全生命共同祖先的性质与基因组及在这个根上不同的生物域的形成。

安东尼奥·拉兹卡诺尽管在过去的几个月里需要照顾家人的健康，但仍然不懈致力于对本书章节内容的平衡和润色。多位作者创作出他们领域中最富有思想性、权威性和可读性的文章，令读者可以把握这个领域研究主题的背景、目的和进展。我很感谢薛红、宓位建、唐家乐、吴肇坚及陈建欢的研究合作及对整理本书的协助。内子梁绮霞的理解和支持对这本书的完成也至关重要。

王子晖

目　录

1 引言

王子晖

1.1 古老的话题，年轻的学科

绝大多数人类文明都考虑过生命起源的问题并试图找出答案，科学的文明更是如此，而对合理答案的探求已经经历了几个世纪。

直到 17 世纪末期，用自然产生解释非同类繁殖是无可置疑的，其典型例证就是如吸虫等寄生虫在人体内的繁殖。Rene Descartes 的机械原理巩固了这个观点，他将自然科学简化为物质和运动，认为搅动分解的物质就可以产生生命，但是 Francesco Redi 用实验证明，蛆并不会从腐败的肉中产生，而是因为苍蝇产卵。同样，Georges Cuvier 也发现自然产生是不可接受的，并宣称“生命总是由生命而来，它可以被转化，但不能被创造”。最后，当 Louis Pasteur 证明通过加热杀菌可以停止由细菌引起的腐败现象时，自然产生说就偃旗息鼓了。

当然，能够证明蠕虫和细菌在几天或几年内不是自然产生的，并不等于否认生命最初是由自然产生的可能。早期地球上的生命起源显然是另一种意义上的自然产生，它是笛卡儿哲学更合理的表达，将物质具体为原子和分子，而将运动具体为进行化学进化所必需的动能。1828 年，F. Wohler 用氰酸盐和铵盐成功合成了尿素这种人体内含有的成分，打破了无机物和有机物在概念上的屏障。20 世纪初期，Archibald Macallum 提出，生命来源于一种微粒状物质，“当我们试图解释生命起源这个问题的时候，我们不需要假设出一种高度复杂的有机体……来作为所有生命的最原始祖先，而更应该选择由少量分子组成、小到超越了最强显微镜极限的那些”，同时，它会经历无尽的分子结合过程，这样“正确的结构会赋予超显微粒子生命体的化学特性”。

这样，为了产生生命，就有了前生物进化和自然选择。之后，W. Lob 通过将湿润的甲酰胺暴露于静电或紫外线，合成了单种氨基酸，如甘氨酸；而 E. C. Baly 用紫外线作用于二氧化碳和水，合成了甲醛和糖类。1924 年，A. Oparin 明确提出了一个前生物的可能方案，其中来自无机物的有机物首先被组织到聚析液小滴（coazervate droplet）中；“一旦这些小滴与周围介质有了明确边界，它们马上就会具有一定程度的独立性……在增长速度上相互竞争”，最后“就产生了最简单的原始生命体”。JBS Haldane 坚决认为，有机物的前生物合成“必定会一直积累，直到原始海洋达到稀汤的浓度”，这就是传说中“原生汤”（primordial

soup）的由来[6~8]。

受 Harold Urey 提出的还原性大气在前生物合成中重要性的启发，1953 年，斯坦利·米勒（Stanley Miller）设计并进行了前生物合成实验，用电来击穿由甲烷、氨、氢气和水模拟的原始大气，这些气体触电产生了氨基酸，他的结果震动了世界[9]。在关于自然产生说的科学辩论出现三个世纪后，随着生物化学、分子生物学、遗传学和太空探索的发展，前生物进化成为科学前沿。21 世纪的基础研究中，可能只有少数的科研领域会像前生物进化和太空生物学一样获得丰硕的研究成果。

1.2 太空生物学领域

1.2.1 对水的需求

在太空生物学探索中，对水的探求是始终不变的[10]。它的重要性体现在许多方面。

第一，水有非常好的物理性质。它在冰点的异常膨胀及高比热容，使得水很难完全结冰。而较高的介电常数则令水成为带电离子和氢键的良好溶剂[11]。若没有离子和氢键，通过质子梯度产能及模板介导的复制过程将不能进行。蛋白质、核酸和类似的多聚物也需要水作为溶剂，脱水后，活细胞会坍缩失活，并常常立刻死亡。

第二，大气中的二氧化碳（CO_2）溶解于水，形成碳酸，它能够将硅酸盐岩石（主要是 $CaSiO_3$）风化为碳酸钙（$CaCO_3$ 或白垩），这个过程可以减少 CO_2，并将它以不溶的碳酸钙形式储存在海底。在 4 亿年内，地球大气内的所有 CO_2 都将被去除。最主要的再生过程“碳硅循环”是通过板块构造运动进行的：海底的碳酸钙沉入地幔，在那里被加热并分解，通过火山口将 CO_2 重新释放到大气中。若没有这种依赖于液态水的再生过程，地球上游离 CO_2 的耗尽将导致所有依赖于 CO_2 的自养活动及大气有机物合成终止，而后者正是异养养分的主要来源[12]。

第三，亲水物质易溶于水，而疏水物质如油、脂肪和石油则易与水分离，它们的不相溶性是地球上细胞膜构建的基础。当这两类化合物搅到一起时，它们不会混合，而会立刻分成两个相态，例如海上的浮油，尽管有巨浪搅动，但仍漂浮在水面上。这两类物质的不相溶性是细胞膜构建的首要条件。膜的主要组分磷脂为两亲化合物，该结构的头部亲水，尾端疏水。悬浮在水中时，它们会自动排列成双层，两层的尾部都在中间，远离水，并使两层间的疏水-疏水接触最大化。两层的亲水头部都露在外面，被水分子包围。这些双层膜使所有细胞都具有独立性，这对细胞水平上的自然选择是必需的。没有水就不能形成细胞膜。一些化合

物，如乙醇、二甲基甲酰胺和二甲亚砜，即使具有一定的水溶性，但脂溶性太高，不能够形成稳定的膜。没有膜，就没有自然选择，那样怎么产生生命呢？这就是为什么2008年NASA的凤凰号火星登陆器在火星土壤中确定了水的存在是如此重大的突破的原因。

1.2.2 可居住区

对于任何地外区域是否适合太空生物学的探索，一个基本概念是可居住区（Habitable Zone，HZ），也称为Goldilock Zone，它不太冷也不太热。组成地球生命体的主要化学元素为氢、碳、氮、氧（H、C、N、O）及磷、硫、铁（P、S、Fe）。H、C、N、O在元素周期表中都具有较低的原子序数，它们可以与其他原子形成单键（H）或多键（C、N、O），从而形成生物多聚体。用更大的原子构成生物多聚体将会产生大分子，它们的可能构型在空间上受到很大限制。考虑到地球生命体所需的蛋白质、酶、核酸和核糖体构型的适当折叠，在任何行星或卫星上，这些较轻的H、C、N、O元素都更有可能构成生物分子的核心。但是，H、C、N、O构成的多键分子通常不能耐受极高的温度。从另一个角度说，所有化学反应都会在低温下变慢，每降低10℃，反应速度会下降为原来的1/2～1/3，这种下降主要是因为反应物分子在低温状态下能量较低，不易跨过能垒使化学反应发生，这对宇宙中的所有行星（及其卫星）都适用。因此，温度太高或太低都是难以产生生命的。水的冰点和沸点将可居住的范围缩小到－5～110℃，这亦是地球上嗜极微生物的耐受极限。

因此，基于可居住性需要行星表面具有液态水这个前提，对于主序星，其具有$CO_2/H_2O/N_2$大气的主序星类地行星HZ的内缘由光解和氢逃逸造成的水分缺失来决定。HZ的外缘由CO_2云的形成决定，通过增加反照率（对太阳/恒星能量的反射）来降低行星表面温度。太阳系中对这两个距离的保守估计分别是0.95～1.37 AU（天文单位，1AU＝1496×10^8km，为地球到太阳的平均距离）。由于太阳光度随年龄增加，HZ也随时间扩大。对太阳系中一直都是可居住区范围的保守估计为0.95～1.15 AU，但其他因素也会影响HZ[10,12,13]。

（1）对地球来说，一方面，如果太阳不可预料地变暗，CO_2将浓缩成为CO_2云，增加反照率并降低地表温度，这将带来更多的雪和冰，使反照率更高，导致“失控冰川效应”；另一方面，若CO_2的极度排放增加温室效应，或太阳辐射强度不可预料地增加，那么地表温度将会升高，向大气中蒸发更多的水并减缓对空间的红外辐射，导致“失控温室效应”。

（2）恒星影响其行星的潮汐，潮汐阻碍会减缓行星的自转。在一个潮锁半径（tidal lock radius）内，自转会与公转同步。这种情况下，行星的一面会一直对着恒星，并持续不断地加热，另一面则始终是黑暗。由于亮的那面持续加热，水

可能会被完全蒸发，即使这颗行星两面的平均温度并不极端。

（3）重元素（如 Fe），可以进行与能量生成偶联的可逆氧化还原反应，因此，一颗行星应该需要一些重原子。

（4）在地球年龄约为 5 亿年时，与一颗大小相当的天体碰撞产生了月球，但从那时起，只发生过与较小行星体的碰撞。要在足够长的时间内都具有可居住性，一颗行星需要处在不会发生频繁大碰撞的星河区域中。

（5）由于碳硅循环不仅依赖于水，也依赖于板块运动，所以一定的板块构造运动也是必需的，但过度的板块运动也会破坏生命的稳定。

（6）行星或卫星如果太小，如月球，将没有足够强的引力去维持水或大气；行星表面积如果太大，如木星，就会被广阔的液氢海洋所覆盖。

（7）主序星的寿命随质量下降。对超过太阳质量 2.2 倍的恒星，其中心氢燃烧的持续时间不会超过 8 亿年，根据地球的经验，这段时间可能足以产生生命，但不能产生智能生命。

根据可居住区的这些必要条件，宇宙中生命的出现应该不会很普遍。但是，对地外生命的探索是如此独特，太空生物学正更加努力，将目标集中于第 3～5 章中提到的太阳系内一些最有可能的位置。最后，随着越来越多的太阳系外行星被发现，它们有很多都落在其主星的 HZ 内，这也增加了发现地外生命的可能性。

1.2.3 泛种论

早期地球是前生物进化起始的合适位置，但并不一定是唯一位置。19 世纪，H. E. Richter 提出了石泛种论，认为陨石将生殖细胞从空间带到地球。20 世纪早期，Svante Arrhenius 提出了辐射泛种论，认为微生物可能被行星喷射出来，在星际中散播。通过恒星的辐射压力，它们可能到达地球，带来了生命的种子。P. Becquerel 研究了高度真空和极低温度下孢子和细菌与紫外线的作用，得出的结论是微生物会在到达地球的途中被毁灭。

近年来，在太空实验室和人造卫星（LDEF、EURECA、FOTON）上进行的实验证实，太阳紫外线是对无遮蔽的孢子最致命的因素，但如果能在岩石内部受到保护，孢子就能在太阳紫外线和大量宇宙射线中存活[7,14]，虽然受岩石保护可能也意味着会难以从岩石中脱离出来。

由于胚种论似乎不那么可信，科学探索仍应集中在地球的前生物进化上。而且，任何可能是胚种源起的行星都一定和早期地球一样，具有适合于它产生然后送出生命的条件。因此，“那里”的条件不会与早期地球大相径庭，而对早期地球前生物进化研究得到的成果应该也对“那里”的前生物进化适用。

1.3 到细胞的途径

1.3.1 一些基本问题

在描述活细胞产生的基本轮廓之前，前生物进化研究需要设法解决几个基本问题。有些问题在本书的其他章节讨论，本章将关注以下问题。

（1）是否在生命刚刚开始的时候催化作用就是必需的？

（2）前生物进化需要什么样的能量？

（3）膜、复制和新陈代谢，哪个是最先出现的？

（4）异养和自养，哪个最先出现？

1.3.2 催化作用

生命行使功能需要水环境和适宜的温度，但这些通常不是有机合成的理想条件，因此，绝大多数生化反应在细胞中需被酶催化，或在较早的时候被核酶及其类似物催化。但是，一个悬而未决的问题是，若没有催化作用，前生物进化能否开始。据估计，一个 RNA 基因组的非催化复制时间约为 4×10^8 年，为核酶催化的 10^{10} 倍。因此，在地球早期的 10^9 年内只能进行 2～3 次非催化复制，以这样的速率是不能进行有意义的生物进化的。而且，只有在 RNA 基因组实际寿命大于其复制时间的条件下，才有可能进行扩增，有关的稳定性定律为[15]：

$$kLT < 1$$

式中，k 为 RNA 中磷酸二酯键水解的速率常数；L 为基因组中磷酸二酯键的数量；T 为复制时间。对于一个具有 3 个 50 碱基 RNA 基因的早期核糖有机物而言，L 为 147，而 pH 为 7 时，k 为 1.5×10^{-9}/min，这样 T 必须小于 8.6 年。因此，除非在 8.6 年内完成复制，不然核糖有机物随时可能消失。要这么快完成复制，必须一开始就有催化作用。无论在太阳系内或太阳系外，稳定性定律带来的限制对 RNA 基因组外的其他基因组也同样适用。

1.3.3 能量

化学反应的平衡位置由反应的动力学决定：

$$\Delta G^{0'} = -RT\ln K$$

式中，$\Delta G^{0'}$ 为 pH 7 时的标准自由能；R 为气体常数；T 为温度；K 为平衡常数。若 $\Delta G^{0'}$ 为负值，反应放能，平衡位置偏向于反应物生成产物的方向。若 $\Delta G^{0'}$ 为正值，反应吸能，平衡位置不偏向产物的生成，而是偏向于逆向反应。任何前生物或生物生化系统都依赖于多种吸能反应（及逆浓度梯度运输）。一种有效进行吸能反应的方法是将它与放能反应偶联，这样偶联后的反应就具有一个

合适的 K 值，汽车爬山这个吸能过程与汽油燃烧这个放能过程就是这样偶联的。例如，己糖激酶催化以下反应：

$$葡萄糖 + ATP \rightarrow 葡萄糖\text{-}6\text{-}磷酸 + ADP$$

式中，ATP 到 ADP 的放能水解（－7.3 kcal/mol）与葡萄糖向葡萄糖-6-磷酸转化的吸能反应（＋3.3 kcal/mol）偶联。由于 ATP 水解放出的能量超过了合成葡萄糖-6-磷酸需要的能量，偶联反应能够顺利进行。类似地，生物体内 ATP 可以催化驱动 RNA 合成、DNA 合成、蛋白质合成、核苷酸、氨基酸、多糖的生物合成、激素分泌、肌肉收缩、神经传导等过程。消耗的 ATP 反过来通过糖酵解和线粒体磷酸化再生。在现在的细胞中，ATP 是主要的能量通货或能量载体，其他能够通过水解释放大量能量的高能化合物也可能取代 ATP 的位置，如肌肉中利用磷酸肌酸作为能量储备，在脂肪酸生物合成中利用硫酯。在生物界利用 ATP 作为标准能量通货之前，无机多磷酸盐和硫酯[1,16,17]可能是前生物世界中重要的能量载体。

对于大肠杆菌，每个细胞重 2.8×10^{-13} g，其 DNA 含 4.6×10^{6} 个碱基对，细胞复制一次需要 1.7×10^{10} 个 ATP 分子，也就是每个 DNA 碱基对复制需要 3700 个 ATP 分子[15]。由于对每个碱基对构建 DNA 双螺旋只消耗两个核苷三磷酸，因此一个活细胞的运转消耗远大于构建一个核酸分子。一个晚期的前生物生命基因组最少含有 150 个基因（见 2.7 节），或约 50 000 个碱基对，相当于一个大肠杆菌基因组的 1%。由于前生物生命对能量的利用率可能低于大肠杆菌，前生物生命复制一次可能至少需要 1.7×10^{8} 个 ATP 或硫酯或多磷酸盐分子。高耗能也显示了前生物进化中能源的重要性。

汽油燃烧是一个氧化反应，食物在我们体内转化为二氧化碳也是同样的过程。在现代的细胞中，氧化还原反应通常是最重要的驱动能量载体生成的反应，在早期生命中应该也是这样。在早期地球上，有多种早期电子供体和受体可以分别作为还原剂和氧化剂，为氧化还原反应供能[18]。

（1）早期电子供体：H_2、H_2S、S^0、Fe^{2+}、CH_4、$(NH_4)^+$、CH_2O、光电子。

（2）早期电子受体：CO_2、CO、S^0、NO、$(SO_4)^{2-}$。

其中，光电子由光合作用产生，CO 由大气反应产生，NO 则是通过闪电产生，其他的都可以从火山口或热液口得到。主要的电子受体为 CO_2 和硫，一种非常稳定的电子供体为 H_2，它是由热液口（hydrothermal vents）处橄榄石（$Mg_{1.8}Fe_{0.2}SiO_4$）向蛇纹石（$Mg_{2.7}Fe_{0.3}Si_2O_5[OH]_4$）转化的过程中产生的，这是一种称为蛇纹石化作用的地化作用。环境中多种氧化还原反应都可以参与生物分子的非生物合成。随后，它们还可以参与与能量载体生成相关的胞内电子传递。

1.3.4 膜、复制和代谢，哪个先出现?

除了催化作用和能源这两个必要条件，以下三种成分对于生命也是必需的：膜提供独立的结构单元，复制因子使信息装置得以表达，而代谢作用包含了催化等化学变化。这些要素必须整合到一起产生一个细胞。问题是，它们中的哪一个是最先出现的——生命是通过膜、复制，还是代谢途径产生的呢?

这三个部分中，膜的作用较为直观。它的起始可能不太需要复制或代谢作用的帮助。由陨石[21]或水相 Fischer-Tropsch 反应[22]产生的两亲性分子可以自发形成具有生长和分裂能力的膜囊泡（见 12.4 节）。囊泡能够促进 RNA 多聚作用并维持代谢物浓度，因此它们在最初的作用就很明显了，它们能够为复制和代谢提供独立空间，使在单个“细胞”水平上的自然选择得以发生。在转移到膜囊泡中之前，这些复制因子和代谢物可能会被矿物质和黏土表层吸收，或被限制在矿物微单元中[23,24,17]，但这种条件下的进化在性质上可能比在细胞内更为普遍。

目前的争论在复制和代谢的先导性上游移不定。从复制优先的观点看，核酶及其互补链可以自我复制并通过突变进行新的代谢反应。相比而言，代谢似乎并不具备进化的能力，代谢循环中的要素是否能够自我组织是值得怀疑的，而自动催化循环通常都有内在的困难[25,26]。因此，由于代谢自行进化的这些问题，复制应该在代谢之前。而从代谢优先的观点看，一套燃烧热很低的小分子水溶性有机物就构成了三羧酸（TCA）循环中所有的有机酸，这个发现是非常有启发性的，它意味着这些有机酸可能自发组织成了还原性的 TCA 循环。这样，前生物进化的第一步就可能是通过分子间的联系进行的[27]。相反，若 RNA 复制子随机产生序列，那么这些序列可能没有一点用处。据估计，一个具有 40～60 个核苷酸的三茎环结构可能具有核酶复制酶功能。一个完整的文库由 10^{24} 个这种 40 聚物（40-mer）构成，每个 40 聚物单一拷贝的完整文库约重 1kg。如果我们要找的这种 40 聚物需要两个或更多拷贝，增加的 RNA 需求量将与地球的质量相等[28]。因此，若果真是自行进化，这些复制子将由于随机性彻底陷入困境。在这种情况下，基于分子间联系的代谢过程就需要优先产生，为复制子确定一个方向。

不过，这两种观点的正确性都体现在对相反观点缺陷的论述上，正如 Robinson 的论断：

“但是这两个模型——通常被称为‘基因优先论’和‘代谢优先论’——都仍有显著的问题，并且没有哪一个是完善的或具有明显优势”[29]。

若代谢不能有效引导复制，复制又不能有效引导代谢，那么这个两难处境将只有一种解决方式：它们必须从一开始就共同作用。任何一种合作要想成功，每一方必须对另一方有所帮助。复制子可以通过引发催化新的代谢反应参与代谢。

而代谢是如何帮助复制的呢？由于复制面临的最大挑战是需要从无数没用的序列中挑出有意义的那些，代谢反应能够对此提供帮助。复制子可以帮助产生新的催化代谢类型，而代谢作用可以帮助选择有用的复制子，它们的共同作用便可以迅速发展出一个基于可用结构单元和能源的生命策略蓝图，在不违背基本热力学规律和水溶性的前提下，建立由代谢中固有的分子逻辑构成的代谢通路、循环和网络。

1.3.5 代谢作用对复制子的选择

对复制子的代谢选择有两种机制，分别为代谢物引起的复制子诱导(REIM) 及稳定化引起的复制子扩增 (REAS)。

对 REIM，代谢作用的存在诱导复制子合成。许多 RNA 分子为核酸适体(aptamer)，能与匹配配体特异性结合。图 1.1A 中给出了能与 ATP 结合的 36 碱基的核酸适体的序列。当这种核酸适体与 ATP 结合时，通过与 G8 和 A12 的碱基堆积和氢键作用，ATP 被夹在核酸适体的 A10 和 G11 之间[30]。核酸适体可作为其互补链——反义核酸适体 (anti-aptamer) 的转录模板，反之亦然。一旦核酸适体与反义核酸适体形成双链，基于半保留复制的进一步转录就将停滞。但是，如果 ATP 进入双螺旋，它与反义核酸适体竞争，结合到核酸适体上，使螺旋松弛，在反义核酸适体上诱导合成新的核酸适体。当这条新的核酸适体被另一个 ATP 分子拉开，这个过程会自动重复，合成这个核酸适体的多个拷贝 (图 1.1B)。代谢作用对功能 RNA 的这种特异性诱导会带来核酸适体的选择性扩增，而无功能的 RNA 则不被扩增。突变后，积累起来的核酶更会促进产生新的核酶，可以将诱导代谢物及其类似物转化成新的代谢物，从而扩展代谢通路和网络。

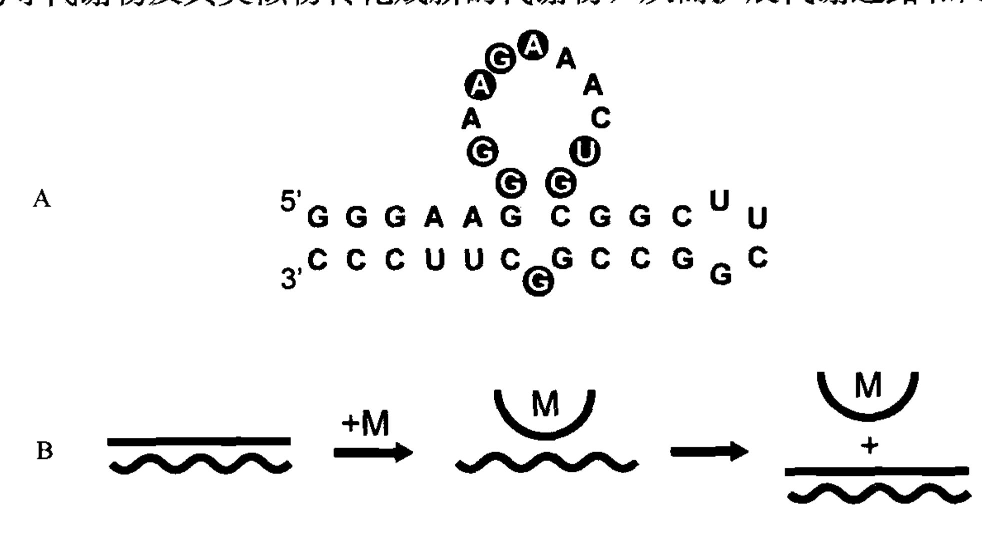


图 1.1 由代谢物引起的复制子诱导 (REIM)。A. 与 ATP 结合的 RNA 适体。用阴影表示保守残基[30]。B. 由代谢物 M 引起的诱导。直线代表适体，波浪线代表反义适体。

De Duve 曾经提出，在化学进化的复制前阶段，各种前体装配成了多种多样的酶，底物与酶的结合能够选择性地稳定这种酶，使之不被水解并增加它的数量，从而可能使它进一步被修饰，并产生新的酶催化反应[31]。酶可以被其底物保护不被降解这个事实，支持了这种机制，但是特异性序列装配成酶带来的困难可能是个限制因素。事实上，生物多聚物稳定化在复制过程中效力更强。通过印记实验，我们知道配体结合，如抑制子结合 DNA，可以保护 DNA 上的相应区域。因此，在 REAS 中，代谢物可以结合匹配 RNA 核酸适体或核酶，从而保护其不被降解。由于 RNA 比蛋白质更易水解，这样的保护作用更可以选择性地保留匹配 RNA 序列，而其他序列则会迅速被水解去除。更重要的是，与核酶一起的稳定效应超过了催化积累。由于催化复制子具有核酶性质，稳定化带来了复制子的积累，可以通过复制进行扩增，并能通过突变产生新的酶促反应，从而扩展代谢作用。图 1.2 显示了 REAS 中这种扩增和突变的结合，它可能用于产生如转化 IMP 到 GMP 的核酶。由于 REAS 是通过无功能复制子的降解来选择功能性复制子，它的效率可能比 REIM 低。但是，因为 REIM 和 REAS 的机制不同，它们在反应时可以进行互补。

$$R_i \xrightarrow{I} R_i\cdot I \rightarrow 反义R_i \rightarrow \cdots \rightarrow 反义R_g \rightarrow R_g \xrightarrow{I} R_g\cdot I \rightarrow R_g\cdot G \xrightarrow{G} R_g$$

图 1.2 通过稳定化作用引起的复制子扩增（REAS）。①IMP（I）结合到适体 R_i，而稳定 R_i；②R_i 直接形成反义 R_i，反义 R_i 突变为反义 R_g。③反义 R_g 转录为核酶 R_g，将 I 转化为 GMP（G）。

复制和代谢作用的有效合作是前生物进化的关键。若没有代谢物的引导作用，复制子不太可能克服其随机序列带来的困境，实现自我进化，这意味着若没有 REIM 或 REAS 这样的选择机制，生命就不太可能产生。

1.3.6 早期复制子

从对 RNA 核酶活性的探索开始，RNA 就超越之前对蛋白质及类蛋白质相应作用的假设，成了前生物复制子中最重要的候选者。对 RNA 是前生物复制子的证据如下。

（1）核酶可以同时作为复制子和催化剂，这解决了先有复制子还是先有生物催化作用的问题。

（2）核糖体多肽合成是由核酶催化的。

（3）因为 RNA 是通过互补碱基配对进行复制，而 RNA 核酸适体能够结合特异性配体，RNA 复制子与 REIM 非常相符。

(4) RNA 中的磷酸二酯键容易水解，因而 RNA 复制子与 REAS 也非常相符。

正如 Orgel 所说，前两条基本上足以证明在今天的 DNA-RNA-蛋白质世界之前，存在过一个 RNA 世界[32]。我们需要确定的是，RNA 复制子功能是从一开始就出现了，还是取代了一些更早的复制子，比如黏土小板[33]、RNA 类似物或肽核酸（PNA）[34]，它能与互补的 DNA 或 RNA 形成双链。我们还不清楚黏土小板是否可以被 REIM 或 REAS 选择。而对 PNA，由于其模板复制方式及更加耐热的骨架，可能符合 REIM，而不符合 REAS。PNA 的另一个很重要的缺陷是缺少能够直接进行前生物合成的 PNA 单体[32]。而相对于 RNA，PNA 的整体适应性较弱，这也意味着它需要更大的分子来实现核酸适体和生物催化活性。第三条和第四条则利于 RNA 和 TNA 这两种复制子的早期利用，TNA 为苏糖核酸，是一种具有苏糖磷酸骨架的 RNA 类似物[35]。而冰[36]和黏土[37,38]能够促进 RNA 多聚化，这个发现也支持 RNA 的早期利用。

关于如何产生细胞还有很多未知的问题，但关键过程的骨架轮廓已经得到。在早期阶段，环境中可用的生物分子与矿物质相互作用，并被黏土催化的代谢作用转化，从而产生 RNA、TNA 或其他与 RNA 类似的复制子。通过 REIM 和 REAS，复制子和代谢物间的相互作用有助于对代谢中有用的核酶复制子（replicator-cum-ribozyme）进行选择，从而扩展代谢反应及生物合成途径。接着复制子与代谢物就会与分布在矿物表面的类脂膜汇合，封装入由这些膜形成的囊泡内。然后会不断发生膜-复制子-代谢物相互作用，逐渐形成第一个细胞。

1.4 异养起源

1.4.1 异养起源说与自养起源说

关于第一个细胞有两种学说，异养起源说认为细胞首先利用环境中可用的有机物，而自养起源说则认为，细胞首先仅仅利用环境中的 CO_2 和单碳化合物来合成其他一切自身化合物（图 1.3）。前生物进化从膜、复制子和代谢物开始，并继续进化到最后一个全生命共同祖先（Last Universal Common Ancestor，LUCA），所有现存生命都由它而来。不管是“自下而上”还是“自上而下”地推断，都得到相同结果[40]。前者由结构单元开始，经过超分子结构和自我延续的自发组织来形成细胞[41~43]，而后者则是确定 LUCA 的性质，然后追溯发展 LUCA 的轨迹。这两种方法都需要知道第一个细胞到底是自养还是异养。

1 期氨基酸和核酸成分的非生物合成实验，以及陨石富含有机物这个发现，催生了 Lazcano 和 Miller 的异养起源说[44]。但是，近年来的重要发现证明，一些

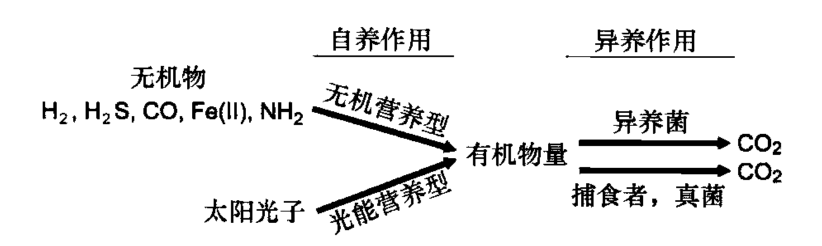


图 1.3 自养和异养生命体（见参考文献 39）。

有机生物分子也可能从热液口及火山口得到：

（1）氨基酸[45]；（2）脂类[22]；（3）糖类[46]；（4）丙酮酸；（5）α-羟基和 α-氨基酸[48]；（6）形成乙酰硫酯的 Wood-Ljungdahl 通路[49]；（7）有机物的 Fischer-Tropsch 合成[50]。

这些发现有利于热液-火山系统内的生命自养起源的可能性，并支持“热起源假说”（Hot Start Hypothesis），即“生命起源于滚烫而富含矿物质的水中，它们从深海热液口喷涌出来”。[51]这导致了一场仍在进行的异养起源对自养起源的争论。

“有机汤（organic soup）理论的支持者认为，大气中通过 Miller-Urey 合成的有机分子或从空间传递到地球的有机分子带来了生命……而表面代谢理论的支持者则认为恰恰相反，声称代谢反应是自养起源的……两边的支持者都对另一方表示了严重的批评，看不到一点妥协的余地”。[17]

对反应产物的定位增加了自养起源说的可能性，而它们在海洋中的稀释则减少了这种可能性[52,53]，但考虑到稀释的程度，也能够“推理出”不那么确定的可能性。现在的生命体，包括异养和自养的，对第一个细胞的代谢性质都不能有所启发。就算确定了 LUCA 是自养的也不能结束争论，因为第一个细胞和 LUCA 可能不具有相同的代谢类型。同样，包含三价铁的“含铁带”的地质形成只能作为前 2.3 Gya* 的证据，只能证明氧化光合作用在早期分布不广，不能排除小范围内的早期光合自养。自养起源和异养起源间的争论仍然悬而未决，也就是说不能确定第一个活细胞的代谢性质。

1.4.2 三个论据的整合

基于遗传密码结构，遗传密码的共进化理论认为，最早环境中只有 10 种 1 期氨基酸可用于蛋白质合成。之后，由于氨基酸生物合成途径的发展及遗传密码的增加[54,55]，产生了另外 10 种 2 期氨基酸，它们中有的是通过 pretran 合成的（图 1.4），如天冬酰胺。这个理论现在已被 pretran 合成的原始起源证明了（见

* 时间单位：1Gya＝10^9 年

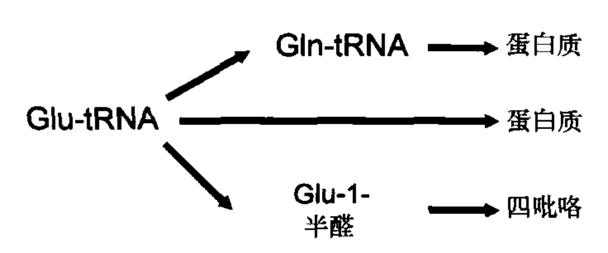


图 1.4 利用 Glu-tRNA 的不同途径。在核糖体蛋白质合成中，中间途径利用 Glu-tRNA 将 Glu 送进蛋白质。上下两条途径显示两种使用 Glu-tRNA 的翻译前合成，即 Gln-tRNA 参与核糖体合成蛋白质的过程，以及在原血红素的生物合成中 Glu-1-半醛作为四吡咯的前体。

参考文献 56 和 14.3 节)，它意味着用于蛋白质合成的氨基酸是由环境中的可利用性决定的，这只可能在细胞异养情况下成立。

大气氨基酸的合成结果证明，环境只提供现今 20 种标准氨基酸的一半。利用高能光线，Kobayashi 等人[57,58]确定了目前最大量非生物合成氨基酸的条件，其产物包括 10 种 1 期氨基酸，而 2 期氨基酸一种也没有。而陨石氨基酸给出了进一步的证明：目前最多样的陨石氨基酸是由 Pizzarello[59]在南极 CR2 号陨石上发现的，也是只有 10 种 1 期氨基酸，没有 2 期氨基酸（见 6.2 节）。

遗传密码结构、大气氨基酸合成和陨石氨基酸这三个独立证据的整合，完全确认了前生物环境中可利用的氨基酸（表 1.1）是什么，并有力证明环境中的可用性是任何一种氨基酸进入早期遗传密码的首要条件。

表 1.1 有关不同氨基酸前生物可利用性的三方面证据的吻合

	Gly	Ala	Ser	Asp	Glu	Val	Leu	Ile	Pro	Thr	Phe	Tyr	Arg	His	Trp	Asn	Gln	Lys	Cys	Met
期	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
辐射合成	+	+	+	+	+	+	+	+	+	+	0	0	0	0	0	0	0	0	n	n
陨石氨基酸	+	+	+	+	+	+	+	+	+	+	0	0	0	0	0	0	0	0	0	0

注：对 1 期和 2 期氨基酸的分类来自参考文献 54。参考文献 57 和 58 描述了由大气辐射生成的氨基酸。文献 59 描述了陨石氨基酸。“+”表示生成或存在，“0”表示不生成或缺乏。“n”表示辐射合成条件下缺硫而无法生产。

如果第一个细胞是自养的，它将会自己合成所有 20 种标准氨基酸，如现在的蓝绿藻，这并没有对表 1.1 中的三个论据给出合理的解释。然而，如果第一个细胞是异养的，这三个论据就刚好是预期的结果，因为细胞如果异养，除了利用环境中的氨基酸，即 10 种 1 期氨基酸，进行蛋白质合成外，别无选择。之后，当细胞氨基酸生物合成途径产生了 2 期氨基酸后，它们也会被利用，以增加蛋白质化学功能。

基于这些，第一个细胞是异养的。

1.4.3 代谢的扩展

异养起源需要环境中具有足够的有机营养分子去启动进化，但并不是说所有最终需要的代谢物都要从环境中获得。例如，即使环境中有磷酸化辅酶 FAD 和

NADH，但它们不能通过类脂膜进入细胞。代谢和生物合成途径需要通过以下机制进行扩展。

(1) 当环境中的一种生物分子被耗尽时，在“逆行演化”中，将产生反向生成这种分子的生物合成途径。通过引入一种催化剂，利用现有的前体催化形成这种生物分子。当这种现有的前体耗尽时，另一个催化步骤就会介入，将前体上游的物质转化为这种前体，诸如此类[60]。比如说，丙氨酸、天冬氨酸和谷氨酸为环境中可用的 1 期氨基酸，若环境中它们被耗尽，就会通过逆行演化进行合成和转化，分别将丙酮酸、草酰乙酸和 α-酮戊二酸转化为丙酮酸、天冬氨酸和谷氨酸。

(2) 在复杂色素（如叶绿素）的生物合成中，这个过程是否能从环境中的叶绿素开始，并反向发展出生物合成途径是很值得怀疑的。相反，这个生物合成途径很可能是通过“渐进扩充”正向进行的，通路中每种中间色素都行使生物色素功能，直到它被通路中渐进扩充产生的高级色素取代。以这种机制，叶绿素 a 的生物合成途径可能是由原卟啉-9 开始，到镁-乙烯基-菲欧-卟琳-a5，最终到叶绿素 a[61]。

(3) 一些二级代谢物（如生物碱和抗生素）的形成，可能是通过“创新性生物合成”进行的，通过催化剂的随机反应合成化合物[54]。由于酶通常没有绝对底物和反应特异性，它们可能会与计划内或计划外的底物进行反应，生成额外的副产物。若一种副产物恰好有用，这个反应就有可能一直进行。例如，即使青霉素对青霉菌十分有用，但它化学性质太不稳定，不能在环境中积累，进行逆行演化。这个生物合成途径的中间物活性也不够，本身不能作为抗生素来激活渐进扩充过程，因此，青霉素的代谢起始可能是通过创新性生物合成进行的。在 1 期氨基酸-tRNA化合物上进行的 2 期氨基酸合成，在 2 期氨基酸遗传编码中起重要作用（见 14.4 节），这是开放性生物合成的一个例子。图 1.4 显示了在许多生命体中通过谷氨酸-tRNA 的谷氨酰胺的翻译前合成，以及可以作为 5-酮戊酸和原血红素前体的 Glu-1-半醛的合成[62]。前生物环境中没有谷氨酰胺（见 9.2 节），而且功能上谷氨酰胺并不比谷氨酸高级。同样，Glu-1-半醛太不稳定，难以在环境中积累，而谷氨酸亦不能替代原血红素的功能，因此，这类生物合成反应不能被称之为逆行演化或渐进扩充，而更像创新性生物合成的结果。

(4) 对于有无机离子参与的催化反应，一个可能适用于多种金属-酶的催化作用构建机制是“酶化无机催化”（enzymatized inorganic catalysis），即一个反应最初由无机化合物催化，之后转化成更有效率的酶促反应（早期则为核酶促反应）。举例来说，无机铁将过氧化氢分解为氧气和水时，催化效率很低。当铁与原卟啉环结合形成原血红素时，其活性增加了 1000 倍。若原血红素与过氧化氢蛋白结合，则活性提高 100 万倍[61]。

1.5 生命的决定性时刻

什么是生命？

如果谁要说出持续时间最长的谜题，“什么是生命”肯定排得很靠前。每个人都有一次生命，人们会珍惜它，且常常为其所吸引。它的迷人之处超越了所有时间、地理和语言的屏障，对整个人类都极其宝贵。对前生物进化而言，绝对有必要解开这个谜题，因为如果不知道生命是什么，就无从知晓前生物阶段何时结束，生命从何开始。一些对生命的定义可以回溯到几十年前[63]。

FG Hopkins，1913——生命最起码的需求是“多相系统中的动力平衡”。

JBS Haldane，1952——一个简单的生命体，如细菌、病毒，包含约 100bt 左右的负熵或信息量，这相当于原始海洋中 10 亿年间自发形成的量。

NW Pirie，1957——“20 年前我认为是不可能有严格定义的。现在人们似乎普遍接受了这一点。”

2002 年，Palyi 等收集了大量定义，现列举如下。

Gustaf Arrhenius——基本因素为自我组织、自我复制通过突变进化、代谢作用和集中封装。

Andre Brack——自我增殖、突变和进化。

David Brin——能量向下流动，秩序、信息和处理能力在狭窄时空束内迅速增加。

David Deamer——半渗透边界结构、多聚催化系统、多聚指令系统。

Christian de Duve——生命就是所有活物的共同点。

Klaus Dose——膜、代谢反应、代谢控制、复制和突变性使得进化成为可能。

Ricardo Guerrero 和 Lynn Margulis——会做出选择、凝聚时光和打破梯度的物质。

Romeu Cardoso Guimaraes——代谢、生长和稳定生殖。

Robert Hazen——多样性的代谢作用和增殖，一系列不连续性显现步骤的概念。

Gerald Joyce——美国宇航局的工作定义为“生命是自我维持的化学系统，并能进行达尔文进化”，其中自我维持即遗传指导的代谢作用。

Vladimir Kompanichenko——能够极力对抗不断进行的自我毁灭过程的组织形式。

Stanley Miller——生命的起源即是进化的起源。达尔文进化需要复制、突变和选择。

Eors Szathmary——Tibor Ganti 的定义是基于代谢网络、大分子模板和封

装膜的，它们都是自动催化的。

Hubert Yockey——具有基因组和遗传密码。

这些定义有本质上的不同，但都认为对生命的定义要基于几种特征的联合，而不仅仅是某一个特征。原因是，极少有单一的特征是生命所独有的。比如，通过在盐的过饱和溶液中加入一点晶体，它在结晶时就会自我增殖/复制，而地球上有许多地化循环比代谢循环更为古老，所有的肥皂泡都有膜，而经过体外进化，核酶也能表现出进化能力。说到底，若进化能力为生命所独有，那么就不会有前生物分子组合能够进化而带来生命了。

除了根据第一个活细胞的实际组分认定的独立特征去定义生命，还有必要指出生命之所以有目标性和独特的精华所在。

David Abel——“我们将'生命'的精华提炼的越多，就越有机会阐明那些已失落的自然发生的途径。”

Hans Kuhn——“有物体在行为中表现出一定的目标性……让我们把这些具有基本新性质（不在前生物世界中能够找到任何原始形式的）的物体称为生命吧”。

而且，也有人提出了要将非生命成为生命那一刻定义出来的这个难题。

Janet Siefert——“我们可以开始假设'生命'出现在地球上那个决定性的时刻或者条件……我认为当以下事件同时发生时，就可以认为生命出现了”，即复制、翻译、控制和细胞壁。

这个定义指出蛋白质酶经翻译生成的重要性，但它排除了 RNA 世界中无翻译的核糖生命体存在的可能，限制性可能太强。翻译并不是一个独立的特征。相反，为了使自然选择有效，通过催化作用表达出来的性状必须完全由复制子内的基因型决定，可以是自身催化，如核酶，也可以是用基因编码生产出酶的替代催化。那些只能复制而不能参与催化的复制子，如扩增中的盐结晶，就根本不能与生命产生任何关系。因此，在地球或其他地方，生命初始时的复制功能不仅需要突变能力，使进化得以发生，也要能够实现催化，要么来源于它们自身（如核酶），要么能够通过翻译生产催化剂（如酶）。因此，Siefert 同时存在论中的翻译应被代谢作用取代，与 Arrhenius、Dose、Ganti 等的定义一致。

膜、代谢和复制这三个原件通过各自独立的前生物起源碰在一起，共同构建了细胞，这仅仅是因为它们能够密切配合。这种配合是通过催化作用、遗传编码、封装、REIM、REAS、能量偶联、转运梯度、信号转导这一整套进化出来的工具巧妙构造的，使分子以比通路和自动催化循环更高复杂性的形式组织起来。组织起来的整体功能远远超出了这三个原件的各自功能，超出的部分具有全新的性质，代表了生命的精华，这是非生命世界不能想象的。这种组织形式通过遗传、变异和进化，具有极强的适应能力，使生命具备以目标性为导向的行为特

征。假以时日，这种突出的适应性就会从单细胞水平发展到多细胞、神经系统和社会水平。基于这一点，生命的决定性时刻就是三个原件成功整合的那个时刻。自我复制是生命的基本特征，并且对进化来说也是首要的，因此，它可以作为一个合适的成功信号。这样生命的决定性时刻可以定义如下。

生命就是膜、代谢作用，以及能够突变并实现催化的复制子整合成的一个整体。这样一个整体的自我复制就是它成功整合的信号，标志着生命诞生的时刻。

因此，当第一次分裂时，活细胞就诞生了。图 1.5 显示了这个历史性的事件。

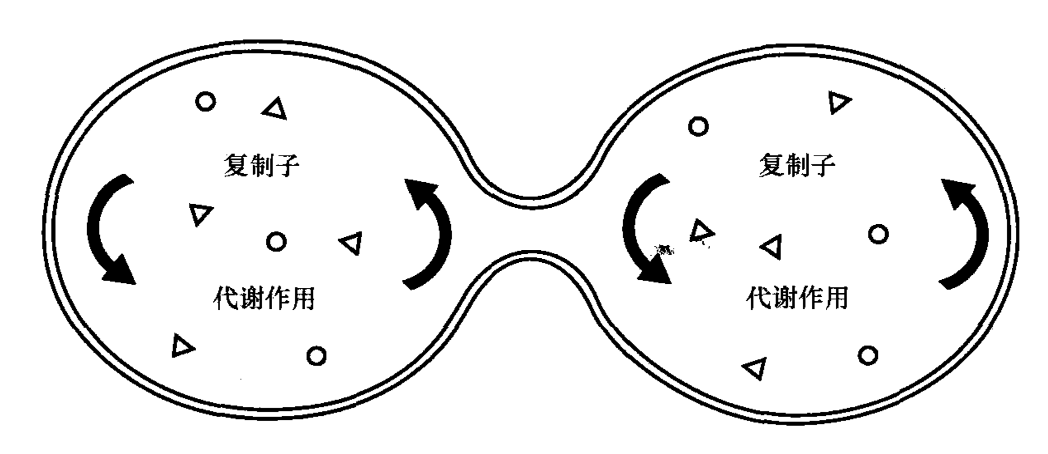


图 1.5　生命的决定性时刻。其中硫酯（三角形）和 ATP（圆形）（见 9.6 节）为第一次细胞分裂供能。箭头表示协同进行。

参考文献

扩展阅读

以下可读性很强的书籍能使读者了解有关前生物进化和研究这个课题的多种科学方法的关键问题。

1. De Duve C. Blueprint for a Cell. Neil Patterson Publishers, 1991.
2. Lahav N. Biogenesis. Theories of Life's Origin. Oxford University Press, 1999.
3. Luisi PL. Emergence of Life. From chemical origins to synthetic biology. Cambridge University Press, 2006.
4. Morowitz HJ. Beginnings of Cellular Life. Yale University Press, 1984.
5. Palyi G, Zucchi C, Caglioti L, eds. Fundamentals of Life. Elsevier, 2002.

专题文献

6. Fairley J. The spontaneous generation controversy from Descartes to Oparin. John Hopkins Univ Press 1977.
7. Raulin-Cerceau F. Historical review of the origin of life and astrobiology. In: Seckbach J, ed. Origins, Genesis, Evolution and Diversity of Life. Kluwer Academic Publishers, 2004:17-33.
8. Oparin AI (translated by Morgulis S). The Origin of Life. MacMillan 1938:109-195.
9. Bada JL, Lazcano A. Prebiotic soup—revisiting the Miller experiment. Science 2003; 300:745-746.
10. Franck S, von Bloh W, Bounama C et al. Habitable zones in extrasolar planetary systems. In: Horneck G, Baumstark-Khan C, eds. Astrobiology. The Quest for the Conditions of Life. Springer-Verlag, 2001:47-55.
11. Henderson LJ. The Fitness of the Environment: an Inquiry into the Biological Significance of the Properties of Matter. Boston: MacMillan, 1913.
12. Ulmschneider P. Intelligent Life in the Universe. Springer Verlag, 2003:51-71.
13. Kasting JF, Whitmire DP, Reynolds RT. Habitable zones around main sequence stars. Icarus 1993; 101:108-128.
14. Horneck G, Mileikowsky C, Melosh HJ et al. Viable transfer of microorganisms in the solar system and beyond. In: Horneck G, Baumstark-Khan C, eds. Astrobiology. The Quest for the Conditions of Life. Springer-Verlag, 2001:57-76.
15. Wong JT, Xue H. Self-perfecting evolution of heteropolymer building blocks and sequences as the basis of life. In: Palyi G, Zucchi C, Caglioti L, eds. Fundamentals of Life. Elsevier, 2002:473-494.
16. Schwartz AW. Phosphorus in prebiotic chemistry. Phil Trans R Soc B 2006; 361:1743-1749.
17. Russell MJ, Martin W. The rocky roots of the acetyl-CoA pathway. Trends in Biochem Sci 2004; 29:358-363.
18. Canfield DE, Rosing MT, Bjerrum C. Early anaerobic metabolisms. Phil Trans R Soc B 2006; 361:1819-1836.
19. Zierenberg RA, Adams MWW, Arp AJ. Life in extreme environments: hydrothermal vents. Proc Natl Acad Sci USA 2000; 97:12961-12962.
20. Oro J, Lazcano A. A holistic precellular organization model. In: Ponnamperuma C, Eirich FR, eds. Prebiological Self-Organization of Matter. A. Deepak, 1990:11-33.
21. Deamer DW. Prebiotic amphiphilic compounds. In: Seckbach J, eds. Origins, Genesis, Evolution and Diversity of Life. Kluwer Academic Publishers, 2004:77-89.
22. Rushdi AI, Simonet BRT. Lipid formation by aqueous Fischer-Tropsch type synthesis over a temperature range of 100-400°C. Orig Life Evol Biosph 2001; 31:103-118.
23. Wachterhauser G. Before enzymes and templates: theory of surface metabolism. Microbiol Rev 1988; 52:452-484.
24. Wachterhauser G. From volcanic origins of chemoautotrophic life to Bacteria, Archaea and Eukarya. Phil Trans R Soc B 2006; 361:1787-1808.

25. Orgel LE. Self-organizing biochemical cycles. Proc Natl Acad Sci USA 2000; 97:12503-12507.
26. Orgel LE. The implausibility of metabolic cycles on the prebiotic Earth. PLoS 2008; 6:e18.
27. Morowitz HJ, Kostelnik JD, Yang J et al. The origin of intermediary metabolism. Proc Natl Acad Sci USA 2000; 97:7704-7708.
28. Joyce GF, Orgel LE. Prospects for understanding the origin of the RNA World. In: Gesteland RF, Cech TR, Atkins JF, eds. The RNA World, 2nd ed. Cold Spring Harbor Laboratory Press, 1999:49-77.
29. Robinson R. Jump-starting a cellular world: investigating the origin of life, from soup to networks. PLoS Biology 2005; 3:e396.
30. Puglisi JD, Willaimson JR. RNA interaction with small ligands and peptides. In: Gesteland RF, Cech TR, Atkins JF, eds. The RNA World, 2nd ed. Cold Spring Harbor Laboratory Press, 1999:403-425.
31. De Duve C. Selection of differential molecular survival: a possible mechanism of early chemical evolution. Proc Natl Acad Sci USA 1987; 84:8253-8256.
32. Orgel LE. Prebiotic chemistry and the origin of the RNA World. Crit Rev Biochem Mol Biol 2004; 39:99-123.
33. Cairns-Smith AG. The Life Puzzle. University Toronto Press, 1971:129-136.
34. Sen S, Nilsson L. MD simulations of homomorphous PNA, DNA and RNA single strands: characterization and comparison of conformations and dynamics. J Am Chem Soc 2001; 123:7414-7422.
35. Schoning K, Scholz P, Guntha S et al. Chemical etiology of nucleic acid structure: the alpha-threofuranosyl-(3'→2') oligonucelotide system. Science 2000; 290:1347-1351.
36. Monard P-A, Kanavarioti A, Deamer DW. Eutectic phase polymerization of activated ribonucleotide mixtures yields quasi-equimolar incorporation of purine and pyrimidine nucelobases. J Am Chem Soc 2003; 125:13734-13740.
37. Ferris JP, Hill AR, Liu R et al. Synthesis of long prebiotic oligomers on mineral surfaces. Nature 1996; 381:59-61.
38. Ferris JP. Montmorillonite-catalysed formation of RNA oligomers: the possible role of catalysis in the origin of life. Phil Trans R Soc B 2006; 361:1777-1786.
39. Nealson KH, Conrad PG. Life: past, present and future. Phil Trans R Soc B 1999; 354:1923-1939.
40. Jortner J. Conditions for the emergence of life on the early Earth: summary and reflections. Phil Trans R Soc B 2006; 361:1877-1891.
41. Maturana H, Varela F. Autopoiesis and Cognition: The Realization of the Living. Springer, 1980.
42. Margulis L, Guerrero R. Origins of life to evolution of microbial communities: a minimalist approach. In: Ponnamperuma C, Eirich FR, eds. Prebiological Self-Organization of Matter. A. Deepak, 1990:261-278.
43. Lehn JM. Toward self-organization and complex matter. Science 2002; 295:2400-2403.
44. Lazcano A, Miller SL. On the origin of metabolic pathways. J Mol Evol 1999; 49:424-431.
45. Hennet RJC, Holm NG, Engel MH. Abiotic synthesis of amino acids under hydrothermal conditions and the origin of life: a perpetual phenomenon? Naturwissenschaften 1992; 79:361-365.
46. Zubay G. Studies on the lead-catalysed synthesis of aldopentoses. Orig Life Evol Biosph 1998; 28:13-26.
47. Cody GD, Boctor NZ, Filley TR et al. Primordial cabonylated iron-sulfur compounds and the synthesis of pyruvate. Science 2000; 289:1337-1340.
48. Huber C, Wachterhauser G. α-Hydroxy and α-amino acids under possible hadean, volcanic origin-of-life conditions. Science 2006; 314:630-632.
49. Martin W, Russell MJ. On the origin of biochemistry at an alkaline hydrothermal vent. Phil Trans R Soc B 2007; 362:1887-1925.
50. McCollom TM, Seewald JS. Abiotic synthesis of organic compounds in deep-sea hydrothermal environments. Chem Rev 2007; 107:382-401.
51. Zimmer C. How and where did life on Earth arise? Science 2005; 309:89.
52. Bada JL, Fegley Jr B, Miller SL et al. Debating evidence for the origin of life on Earth. Science 2007; 315:937-938.
53. Wachtershauser G, Huber C. Response to "Debating evidence for the origin of life on Earth". Science 2007; 315:938-939.
54. Wong JT. Coevolution of the genetic code and amino acid biosynthesis. Trends Biochem Sci 1981; 6:33-35.
55. Wong JT. Coevolution theory of the genetic code at age thirty. BioEssays 2005; 27:416-425.
56. Wong JT. Question 6: Coevolution theory of the genetic code: a proven theory. Orig Life Evol Biosph 2007; 37:403-408.
57. Kobayashi K, Tsuchiya M, Oshima T et al. Abiotic synthesis of amino acids and imidazole by proton irradiation of simulated primitive Earth atmospheres. Orig Life Evol Biosph 1990; 20:99-109.
58. Kobayashi K, Kaneko T, Saito T et al. Amino acid formation in gas mixtures by high energy particle irradiation. Orig Life Evol Biosph 1998; 28:155-165.
59. Pizzarello S. Meteorites and the chemistry that preceded life's origin. In: Wong J T-F, Lazcano A, eds. Prebiotic Evolution and Astrobiology. Austin: Landes Bioscience, 2008.
60. Horowitz NH. On the evolution of biochemical syntheses. Proc Natl Acad Sci USA 1945; 31:153-157.
61. Granick S. Speculations on the origins and evolution of photosynthesis. Annals NY Acad Sci 1957; 69:292-308.
62. Randau L, Schauer S, Ambrogelly A et al. tRNA recognition by glutamy-tRNA reductase. J Biol Chem 2004; 279:34931-34937.
63. Pirie NW. Some assumptions underlying discussion on the origins of life. Annals NY Acad Sci 1957; 69:369-376.
64. Palyi G, Zucchi C, Caglioti L. Short definitions of life. In: Palyi G, Zucchi C, Caglioti L, eds. Fundamentals of Life. Elsevier, 2002:15-55.

2 最小的细胞

Markus Pech，Knud H. Nierhaus

2.1 引言

这一章，我们讨论在细胞得以存在的最一般和最低限度的条件下生命的起源问题，尤其是考虑生命必需的“最小的一套基因”这个问题。

所有的有机体都由细胞组成。每一个细胞都是一个实体，充满各种各样的化学物质和亚细胞结构，比起水溶液，它更像一种溶胶（表 2.1），通过膜或细胞壁与其他细胞区分开来。活细胞与非生命的化学系统最主要的不同点为它们可以：①自养或异养；②复制或增殖；③分化；④产生化学信号；⑤进化。

表 2.1 一个细菌中大致的分子组成（见参考文献 23）

组成	质量百分比/%	每细胞中大约分子数	不同种类数
水	70	4×10^{10}	1
DNA	1	1～4	1
RNA	6	5×10^{5}	3000
蛋白质	15	1×10^{6}	3000
核苷酸及其前体	0.4	1.2×10^{7}	200
氨基酸及其前体	0.4	3×10^{7}	50
碳水化合物（多糖及其前体）	3	2.5×10^{8}	200
脂类	2	2.5×10^{7}	50
离子	1	2.5×10^{8}	20
废物及中间体	0.2	1.5×10^{7}	200

2.2 细胞的化学组成

对活细胞的分析显示其化学组分与无机物有很大不同。细胞不是化学元素的随机组合，而是一个有选择的化学系统，由严格的一套元素组成，主要有氢、碳、氮、氧、磷和硫（图 2.1），它们占细胞重量的 99%以上[1]。值得注意的是，细胞中含量最丰富的分子一点也不特殊——水，占活细胞重量的 70%，这样多数细胞内的反应都在水环境中进行。不同的是，所有细胞结构和生物分子都含有大量的碳。最重要的是，碳不仅可以和许多其他元素结合，也能和其自身结合，

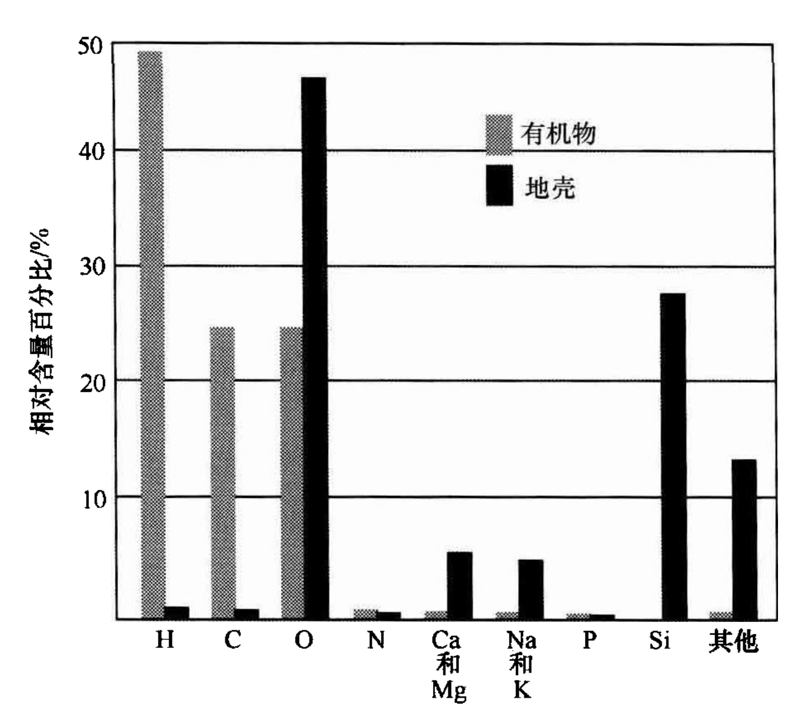


图 2.1 地壳中非生命世界化学元素的相对含量和有机物中的比较。相对含量由所占的原子的百分比表示（见参考文献 25）。

组成链和环，这样产生大量不同的和复杂的化学结构，而且其大小没有明显的上限。

主要的细胞组分是被称为大分子的化学结构，由独立的结构单元通过特殊方式连接构成。这样一个单独的结构单元被称为单体，通常为小的有机组分，由它们的化学性质分类。在细胞生化中有四类作为细胞组分的重要单体：糖类，为构成多糖的单体；脂肪酸，是脂类的单体；核苷酸，是核酸（DNA 和 RNA）的基本结构；氨基酸，是蛋白质的单体。核酸和蛋白质可看成信息大分子，因为它们内部的单体序列是高度特异的，可携带生物学信息及表达这些信息的途径。不同的是，多糖和脂类中单体的序列常常高度重复，并且序列本身通常在功能上不具有重要性。但是，对任何大分子，在化学上，单体的序列和性质在区分它与相关大分子时都起重要作用。

2.3 原核和真核细胞

在生命世界的种系发生中，我们可以区分两种主要的细胞类型——原核细胞和真核细胞，它们在结构上有很大不同（图 2.2）。原核生物被分为两个域：细菌和古菌，一个最重要的区分点是没有“真正的”细胞核［prokaryote，原核生物，来源于古希腊语 pro（前）＋karyon nut 或 kernel 细胞核］及其他任何膜性细胞器，其基因组 DNA，绝大多数情况下为一个单环染色体，定位于细胞中一个形状不规则的区域，称为拟核。除了水，原核细胞中其他的主要组分包括核糖

体、大分子、小的有机分子（主要为大分子的前体）和多种无机离子。细胞外的屏障为质膜，一个可流动的薄层，在细菌中细胞壁由肽聚糖构成。与原核细胞不同，真核细胞[来源于古希腊语 eu（好、真）+karyon nut 或 kernel]具有细胞核，包含其基因组 DNA，并有双层膜保护。除了这些基本特性，我们发现在细胞质中，除了核糖体和大分子，通常还有一些由膜围成的细胞器和单元，如内质网、高尔基体和线粒体，也有过氧化物酶体和溶酶体这样的囊泡。并且，真核细胞具有细胞骨架，为一系列的胞内结构，可在结构上支撑细胞，并帮助其内部组分组织和运输。原核生物中不同物种的细胞组分大致差不多，比较起来，不同真核细胞的组分可以相差很大，这取决于它们的器官或组织。植物细胞组成中还有一个大的中央液泡，在与光合作用相关的组织中，还含有叶绿体。植物的细胞壁由纤维素和蛋白质组成，很多情况下还含有木质素，这与真菌的细胞壁不同，它由几丁质组成。动物细胞没有坚硬的细胞壁，只有细胞膜作为胞外屏障。

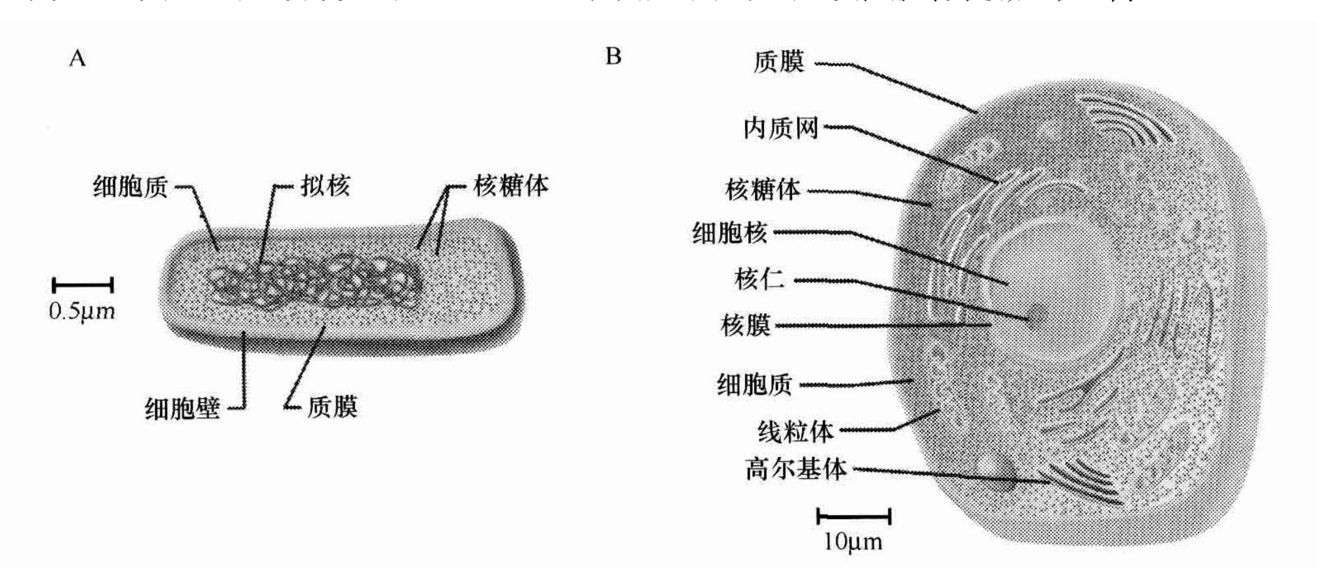


图 2.2　原核细胞（A）和动物真核细胞（B）的比较。解释见正文，来源于参考文献 26（见彩插）。

2.4　地球上的生命起源和遗传密码

生命三个领域的基因组序列分析显示，与能量传递和蛋白质合成、转录、复制相关的主要蛋白质在这三个领域中是一般的和保守的，除了在一些细节上有所不同[2]。

这里，我们已经确定了生命开始之初的必经步骤，通过逆推的方法得到蛋白质合成所需的最少组分，基于我们对大肠杆菌翻译途径的了解，细菌领域是全生命进化树中最原始的（不同观点见 15.2 节）。图 2.3 显示了一个核糖体小亚基 RNA 的全生命树，其上显示了一些最近测定的年代，而对 rRNA、核糖体蛋白质和膜组分比较显示的证据，表明它们在整体上是吻合的。由于蛋白质合成是一个需要能量的过程，我们对早期生命的主要问题，即如何避免能量浪费，进行更

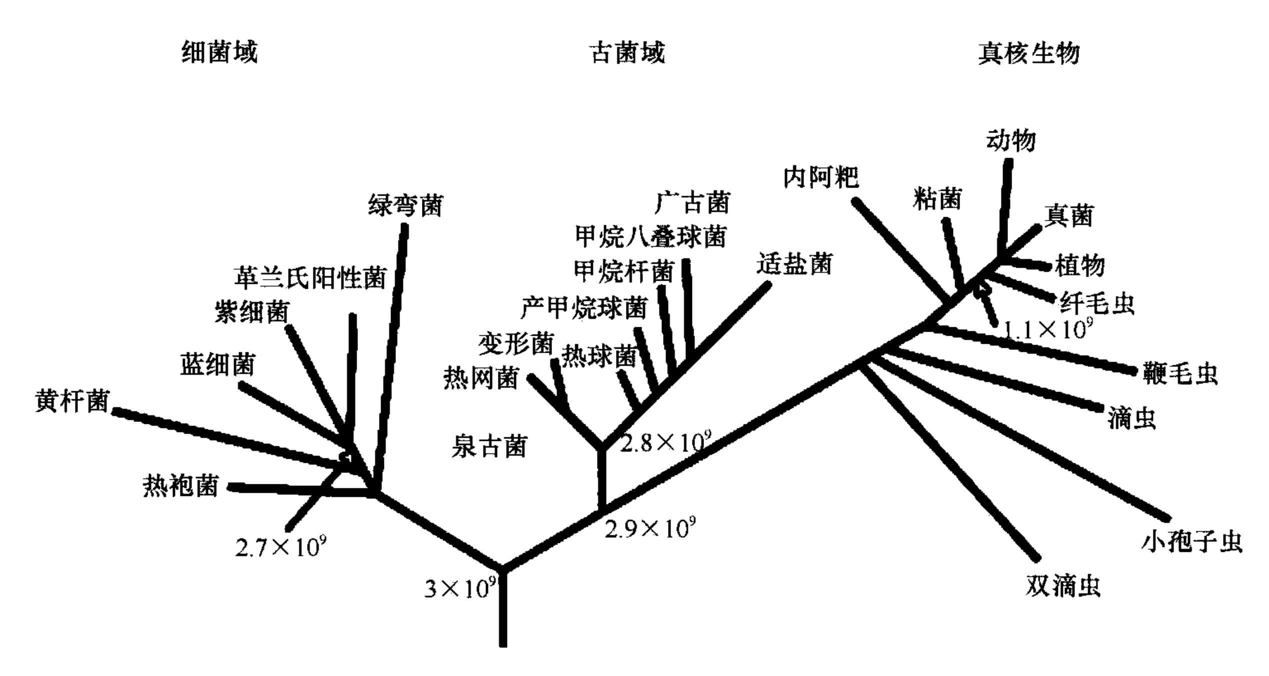


图 2.3 在细菌、古菌和真核生物三个生命领域中小亚基核糖体的种系发生树（见参考文献 6、27、28）（见彩插）。

深一步的探究。

在许多时候，对一些生命特征出现的年代的描述，就像一朵花开的过程，不是一个绝对的厘定。不过，还是可以找到少量可靠的数据。我们对早期生命研究的一个基石是古老细胞的叠层石。有报道的最古老的“叠层石”来自格陵兰 Isua Supracrustal Belt 的岩石，测定年代是 3750 Ma（百万）年以前，但还不确定它们是否是生命最早的印记。有关生命形成的可靠数据还有 3500 Ma 年前古代细胞的残留：①其中之一来自西澳大利亚 Pilbara 区，年代是 3430 Ma 以前，最近的报道支持了这些结构可能是生命起源[3] 的论点；②另一个年代相同且具有微生物生物标记的例子是南非 Baberton 枕状熔岩[4]。含有的古老细胞具有基因（来源于 RNA?）和翻译装置，这表示遗传密码有 32 亿～36 亿年历史。Eigen 及同事的出色的工作[5]，通过来自不同有机体的 tRNA 序列，也估计出遗传密码有（3300±300）Ma 年历史。这表明化学进化、遗传密码的发展和“RNA 世界”的存在，一定挤在 400～800Ma 年间，这符合最早的岩石形成（4000Ma）和最早的细胞存在（3600～3200Ma）之间的时间差距。

另一个重要的里程碑是对沉积在 20 亿年前（2000Ma）地层的大量铁的观察，三价铁沉积物显示了光合作用产物之一——氧气的氧化力。在西澳大利亚，Hamersley 发现了最早的三价铁沉积[6]，在更深的地层中，沉积的通常为二价铁。这表明蓝细菌的光合作用是在 2000Ma 年前产生。由于残留着大气为还原态时的早期生命的特征，细胞内始终是一个还原态的环境，因而大气中污染物氧气的存在对早期生命是一个主要的威胁。带来的结果明显是引起大量的早期生命灭亡。

只有少数细胞由于细胞膜的存在，阻止了大量氧气进入细胞而存活下来。经历了这个长时间的过程，细胞产生了呼吸机制，最终把大气中氧气的存在变成了有利条件。

真核细胞通过吞入呼吸细菌产生了线粒体，特别是一种α-蛋白细菌的祖先[7]，通过利用生命的能量通币——ATP，在能量产生方面带来了巨大的进步。与最古老的产生能量的通路（即厌氧的糖酵解途径）相比，呼吸作用意味着利用能量形成 ATP 的效率提高了 18 倍。我们可以定义与 ATP 产生相关的进化的三个重要阶段（图 2.4）：①呼吸前（>2600 Ma 以前），主要的能量产生途径为厌氧糖酵解，即 C6 化合物（葡萄糖、果糖）分解为 C3 化合物（丙酮酸）。通常认为这些糖类由环境中得来，根据 Stanley Miller 模仿古代大气的组成和物理因子（闪电、温度[8]）的实验，它们通过古代大气中的高温和电击形成。②在 2600～2400 Ma 年前，光合作用产生，使生命不再依赖外源性的高能组分，如 C6 化合物。③在<2000 Ma 年前，一些光合作用形成的酶经过细微的修饰，被再利用，催化呼吸作用的相关反应。实际上，叶绿体中细胞色素 b_6f 复合体所必需的亚基，与线粒体中 bc_1 复合体和细菌的 bc 复合体中的类似，分别将 H^+ 泵到叶绿体间质外、线粒体基质外和细胞质外。呼吸作用的产生在生命进化中是一个主要的进步，使细胞中具有极大的产生能量的潜能，如前述，提高 18 倍。这样，产生多细胞器官和分化成不同组织的道路就此展开。

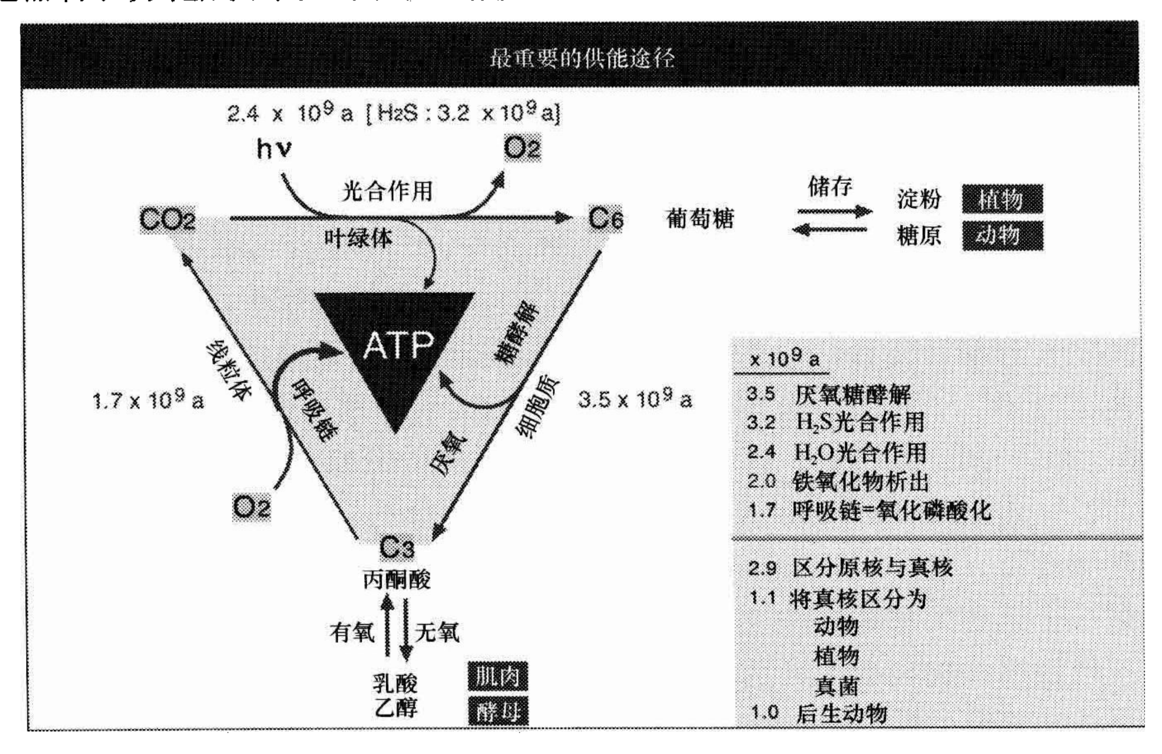


图 2.4　红色显示了三种产生 ATP 的必要途径的年代。ATP 首先由 3500 Ma 年前出现的厌氧糖酵解产生，接着来自光合作用，最后来自呼吸作用，最后一步为产生高度组织的多细胞生命提供了能量。灰色部分给出了一些附加数据（见彩插）。

2.5 简化现存细菌翻译装置的可能性

为了表示蛋白质合成可能需要的最少组分，有必要介绍一下在这个基本且必须的过程中相关的关键物质。所有活的有机体在发育和功能中所需的指令（基因）都包含在基因组 DNA 中。编码蛋白质的基因首先由 RNA 聚合酶进行转录，产生一个互补的信使 RNA（mRNA）。在蛋白质合成的第二步（翻译）中，mRNA 解码，产生由 mRNA 序列决定的特异性多肽。这中间有一些相关组分：首先是核糖体——蛋白质合成的“工厂”。核糖体是一个由蛋白质和 rRNA 组成的巨大的复合体，也是氨基酸装配成多肽链的地点。其他的重要组分有转移 RNA（tRNA，将氨基酸转运到核糖体），以及起始因子（IF）、延长因子（EF）和终止因子（RF）。

在生命早期，mRNA 的翻译仅在 5′端的一个有义密码子处起始，后来则在一个特殊的起始密码子 AUG 处开始，但都没有 5′-不可译区域和 Shine-Dalgarno（SD）序列（核糖体结合位点）。SD 序列与 16S 核糖体 RNA 的 3′端互补，通过碱基配对保证核糖体-mRNA 复合体的有效结合。与通常理解不同的是，在细菌翻译中，平均只有 55%的 5′-不可译序列包含 SD 序列[10]。非解离的 70S 核糖体可以通过定位在 5′端或其附近的密码子 AUG 起始翻译，在 mRNA 上的第 1～3 个核苷酸处优先进行[11]。但实际上，将多聚 U 翻译为多聚 Phe 的过程，通常采用这种在 5′端不含有密码子 AUG 的古老的起始方式。为了描述将现代细菌的翻译装置最小化的可能性，我们假定所有转录物的翻译都在 mRNA 的 5′端或其附近进行，且在多顺反子 mRNA 中，顺反子间区域像编码核糖体蛋白质的 mRNA 一样短，这样 70S 类型的起始就对除第一个外的所有顺反子有效（R. Gupta 与 K. H. Nierhaus，未发表）。通过 30S 亚基进行的标准起始方式将被称为 30S 类型的起始（见参考文献 12、13）。

2.5.1 tRNA 的数目

遗传密码由 64 个密码子组成，其中三个是终止密码子（图 2.5，左侧）。大肠杆菌中，61 个有义密码子通过 42 个不同的 tRNA 进行阅读。注意当两个 tRNA 的反密码子不同时，认为它们是不同的；含相同反密码子，但反密码子之外的序列不同的 tRNA 被认为是同一种，或称同工 tRNA。如果我们认为最常用的密码子编码一个特异性的氨基酸，那么一套 21 个 tRNA 就能满足翻译所有蛋白质的需要——在延伸过程（延长子 tRNA）中 20 个 tRNA 解码 20 个氨基酸的密码子，加上起始 $tRNA_f^{Met}$，它解码 mRNA 编码区开始处的起始密码子 AUG。但是，对于那些可被最少的一套延长子 tRNA 解码的密码子来说，mRNA 中对这些密码子的利用是十分有限的。

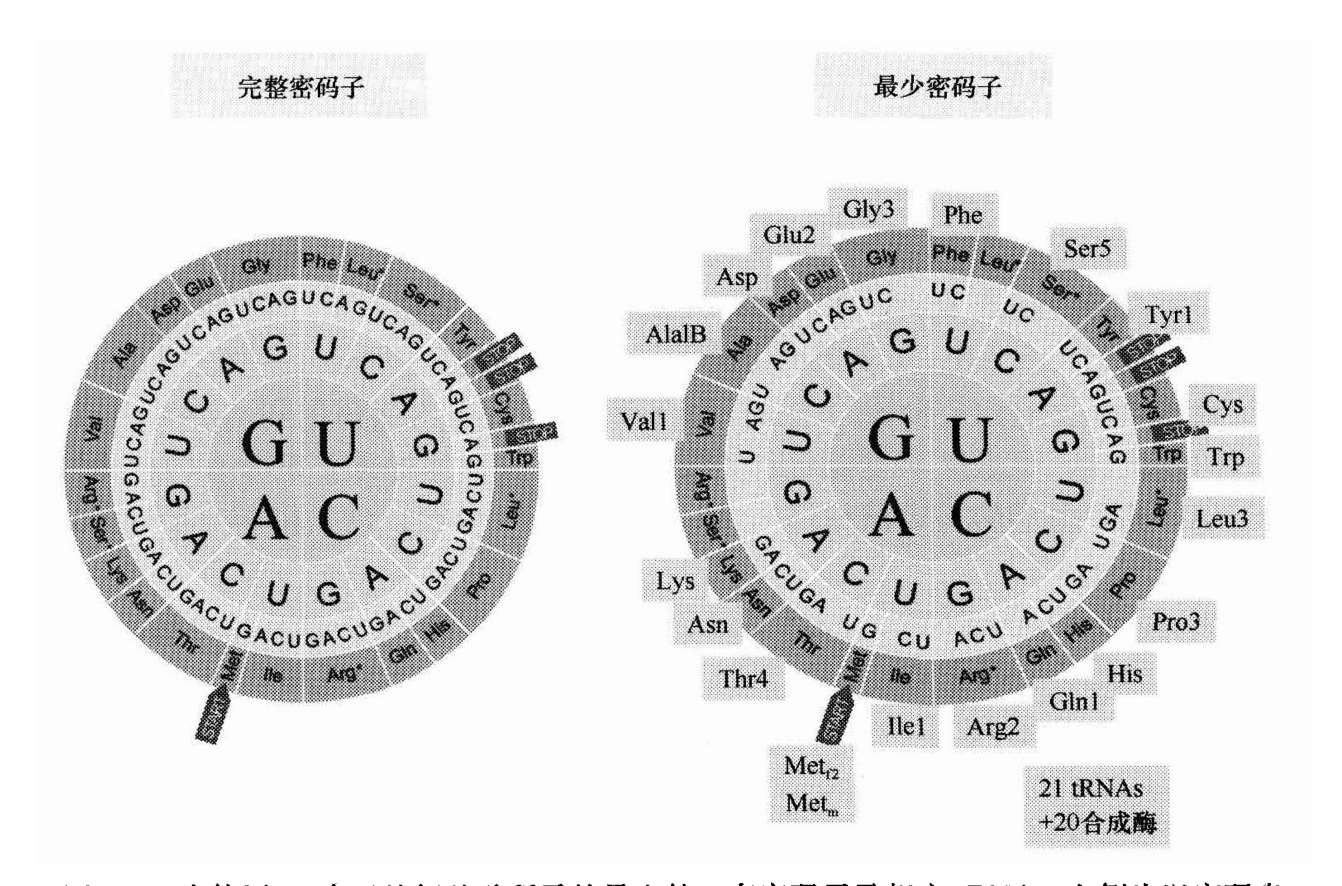


图 2.5　为使用 20 个天然氨基酸所需的最少的一套密码子及相应 tRNA。左侧为以密码盘形式表示的一套标准密码子辞典。右侧为限制性的密码盘，具体解释见文本（见彩插）。

一个反密码子通常可以阅读两个密码子，有时可以读三个，它们在密码子的第三位上不同。例如，密码子 5′-UUC/U 可被反密码子 3′-AAG 阅读。

这种密码子第三位上的“模糊性”称为反密码子的摆动特性。若我们考虑 tRNA 的这种摆动特性，由最小的这一套 20 个延长子 tRNA 翻译的基因组密码子辞典，就被还原表示成图 2.5 右侧中的那种限制性的系统。

2.5.2　核糖体的简化

一个大肠杆菌的核糖体中，并不是所有的分子组成都是必需的。细胞在大约 1/3 的核糖体蛋白质被去除后，仍可保持活性。但是，在很多情况下，缺乏一个核糖体蛋白质的链变得非常脆弱，这可由降低 1/7～1/10 的生长速率看出，或体现在条件致死性上，如低温或高温敏感性（见参考文献 14 和表 2.2）。即使不能确定同时省略这些蛋白质是否会影响细胞活性，也认为一个最小的核糖体应首先定位在这类蛋白质上。有趣的是，已知的最短的 rRNA 存在于高等真核生物的线粒体、核糖体中，比大肠杆菌的 rRNA 短约 30%（低等真核生物，如酵母中的线粒体 rRNA 却类似大肠杆菌 rRNA 的长度）。这些线粒体、核糖体用额外的蛋白质补偿了 rRNA 序列的相对缺乏[15]，因此线粒体、核糖体的总大小比大肠杆

菌的还要大。这样，基于最近对核糖体结构的研究，很难像在 Dabbs 的实验[14]中将核糖体蛋白质的数目减少，同时又将 rRNA 像在线粒体中那样变短。

表 2.2 大肠杆菌中突变引起的核糖体蛋白质缺失

蛋白质	突变名称	表现型
L1	RD19，MV17/10	—
L11	AM68，AM76，AM77	—
L15	AM16-98	cs
L19	AM149	—
L24	AM290	ts
L27	AM125	cs
L28	AM81，AM108	cs
L29	AM111	—
L30	AM10	—
L33	AM90，AM108	cs
S1	VTS03	—
S6	AM80	-未发表
S9	AM83	cs
S13	AM109-113	-未发表
S17	AM111	ts
S20	VT514	ts

注：cs 为冷敏感型，ts 为温度敏感型（见参考文献 24)。

2.5.3 核糖体因子的简化

图 2.3 列出了细菌系统中具活性的翻译因子。翻译起始需三个因子：IF1、IF2、IF3。起始后，翻译和终止由 70S 核糖体控制[16]，反缔合因子 IF3 在前面提过的最小的系统中则不是必需的。

已知延长状态需 4 个因子，经典的 EF-Tu、EF-G 及 EF-Ts，以及最近发现的 EF4[17]（Lep A）（见参考文献 17)。EF4 在 pH$\leqslant$6.5 和（或）高盐（K^+、Mg^{2+}）或低温的生长状况中是必需的，在控制得很好的介质中则不是必需的 (Z. Karim and K. H. Nierhaus，2007，待发表)。

终止状态需 4 个激活因子（包括 EF-G)。RF1 和 RF2 分别解码终止密码子（UAG+UAA）和（UGA+UAA)。当终止密码子是 UAG 和 UAA 时，RF2 可省略，这是由于一些物种缺失 RF2，并用 UGA 编码色氨酸。另一个终止因子——RF3，则通过促进 RF1 和 RF2 从核糖体上释放来调节终止过程。但是，这个因子在大肠杆菌中不是必需的，在最近分析的 191 个基因组中，有 72 个都

不含该因子[18]。核糖体再循环因子（RRF）对翻译终止后 70S 核糖体的解离是必需的，这样 30S 亚基可以再次进入标准的 30S 类型的翻译起始过程。若只利用 70S 类型的起始，如前所述，则 RRF 就不是必需的（见参考文献 16）。

因此，细菌系统中翻译需要的 11 个因子中，只有 6 个在这里的最小系统中是必需的（表 2.3）。

表 2.3　一个最小细胞中蛋白质合成的非必需因子

起始因子	延长因子	终止子
IF1,IF2 and ~~IF3~~	EF-Tu(EF1),EF-G(EF2) and EF-Ts;~~EF4(LepA)~~	RF1,~~RF2~~ and ~~RF3~~ + ~~RRF~~

2.6　对蛋白质合成中能量消耗的优化

如我们之前说过的，新蛋白质的合成是一个急需能量的过程。以下几方面需要能量：①密码子的合成，它是氨基酸的核酸信息单元；②tRNA 通过其聚合酶与正确的（同源的）氨基酸结合的反应；③之后将这个氨基酸加入新生的多肽链的过程，总共需要 10 个高能键。在本文中，一个高能键意味着一个 ATP 或 GTP 中磷酸基团邻近的酸酐键，每一个含约 $\Delta G=-6$ kcal/mol 的能量，这种大量的能量需求解释了为什么细胞产生了复杂的能量消耗控制系统。这完全是离体翻译系统中的现代阿喀琉斯之踵（致命弱点——译者注），在此最多有 5%的能量能真正被用于蛋白质合成，其他的都被不受控制和无用的能流浪费掉了。令人惊异的是，一个细菌细胞可以将其 70%能量消耗于蛋白质合成。

在接下来的部分，我们将概述蛋白质合成中如何避免能量浪费，如何在活细胞中将最多的能量用于有用之处的关键。我们将关注离体系统中能量浪费的两个主要原因：①代谢能量链；②无用的核糖体和过剩的延伸因子 EF-G。

2.6.1　避免代谢能量流失

蛋白质合成中能量（ATP，GTP）再生的标准途径是以 mmol/L 浓度磷酸烯醇式丙酮酸（PEP）和丙酮酸激酶混合，将 PEP 上的磷酸基团转移到 AMP 和 ADP，分别产生 ADP 和 ATP（图 2.6A）。GMP 和 GDP 也可以作为该酶的底物。PEP 含一个磷酸键，是所有生物样本中能量最高的键，含 $\Delta G'^{0}=-12$ kcal/mol。但在这个系统中，由于产生正磷酸盐带来的抑制效应，合成的蛋白质含量是有限的。正磷酸盐副产物不仅降低 pH，也与 Mg^{2+} 离子结合，降低了自由 Mg^{2+} 离子的浓度，严重影响系统中所有核糖核蛋白，尤其是核糖体的结构。Mg^{2+} 的正二价屏蔽了 rRNA 上带负电的磷酸基团，维持了核糖体的高密度。因此自由 Mg^{2+} 浓度的降低增加了磷酸基团间的相互排斥，使核糖体膨胀，降低了效率。

一个防止产生这种不理想的效应的方法是使自由的磷酸基团再循环，通过加

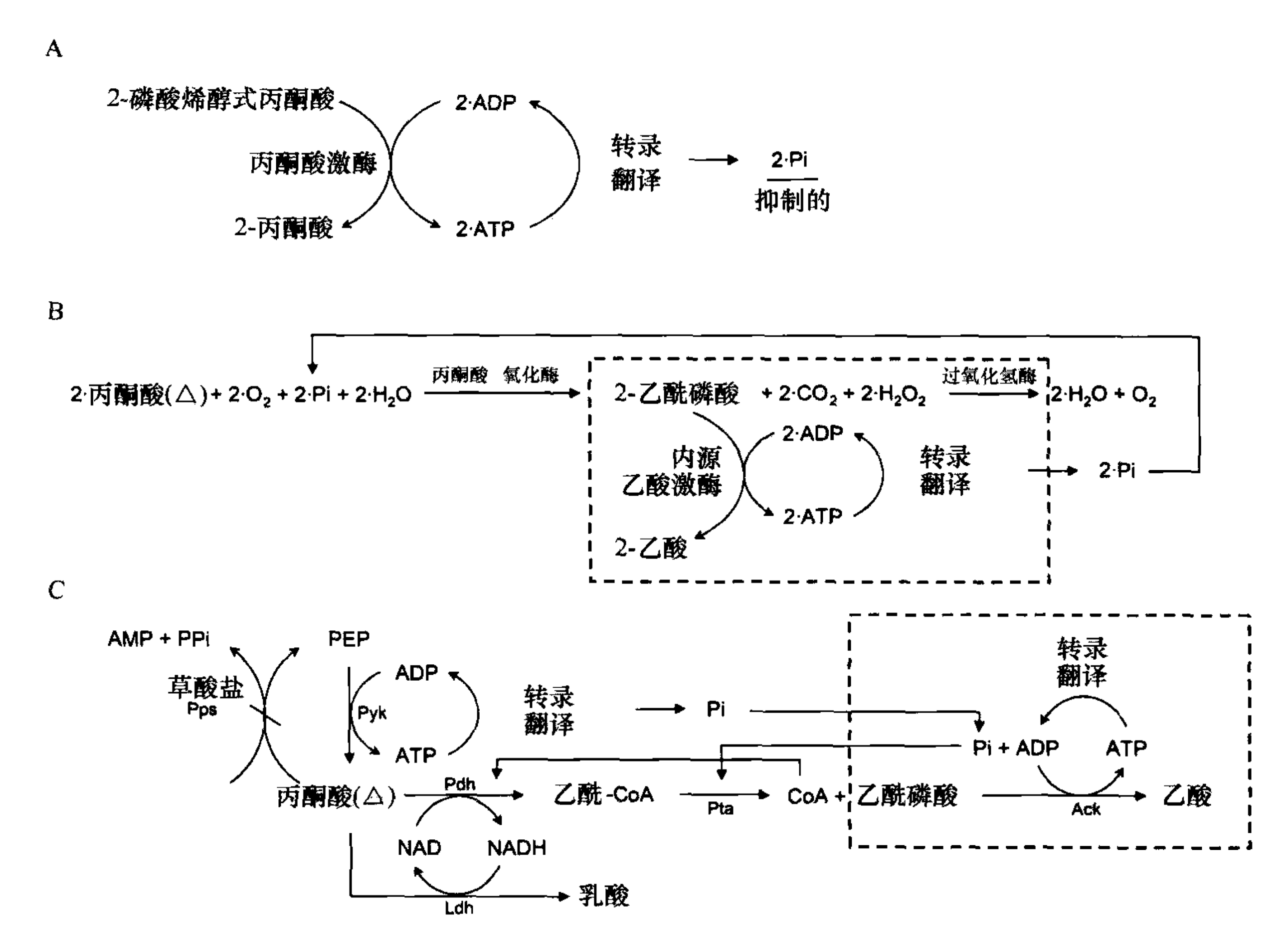


图 2.6 离体蛋白质表达中几种不同的 ATP 再生系统。A. 利用磷酸烯醇式丙酮酸（PEP）和丙酮酸激酶再生 ATP，产生的正磷酸根（Pi）抑制蛋白质合成。B. 改进的能量再生系统，利用丙酮酸、丙酮酸氧化酶和内源乙酸激酶（见参考文献 19）。C. 在近似的活体系统中，由磷酸转乙酰酶（Pta）产生乙酰磷酸。B 和 C 的框中显示了产生 ATP 的一般反应。由于草酸盐的存在，PEP 合成酶（Pps）被抑制。Ack：乙酸激酶；Ldh：乳酸盐脱氢酶；Pdh：丙酮酸脱氢酶；Pyk：丙酮酸激酶（修改自参考文献 29）。

入丙酮酸氧化物和一些辅基（prosthetic groups），如 TPP 和 FAD（分别为维生素 B_1 焦磷酸盐和黄素辅酶Ⅰ）实现。这个系统是一种标准的细胞再生系统，通过利用分子氧和 H_2O（图 2.6B）产生乙酰磷酸，一定情况下这种改进使活体中蛋白质合成明显延长。

但是，这个系统离理想状态仍然很远。对分子氧的需求是主要缺点，因为氧的摄取是受限制的，且当反应量增大时，由于相对表面积减小，这种情况变得更严重。对这个问题，Jewett 和 Swartz 最近发表了一个令人信服的解决方法，将经典的 PEP/丙酮酸激酶系统和由乙酰-CoA 合成乙酰磷酸结合起来，它们在有效地合成乙酰磷酸的同时使正磷酸盐（Pi）再循环（图 2.6C）。这个方法代表了一个突破：第一次能够以每毫升几毫克的反应量合成蛋白质，使离体系统成为在结构和功能分析中合成蛋白质的方法。

2.6.2 避免EF-G过量情况下核糖体的停顿

我们已经提到，翻译的延伸需4种蛋白质，其中，EF-Tu和EF-G是核糖体中心活性延伸循环的组分，可在新生的链上增加一个氨基酸（AA）。EF-Tu将AA-tRNA带到核糖体的解码中心，而EF-G使核糖体上的tRNA易位一个密码子长度（见参考文献20）。这两个因子，与EF-Ts和EF4一起，属于GTP结合（G）蛋白超家族，是重要的调节蛋白，它们与GDP和GTP结合，产生构象变化，该变化由所结合的核苷酸决定。在GTP构象异构体“开启”状态中它们与标靶结合，促进标靶独立反应，为G蛋白结合的步骤提供能量。接下来标靶产生反应已成功进行的信号，在相应的G蛋白酶中心刺激水解反应“GTP→GDP＋Pi”，释放无机磷酸后，G蛋白转化成“关闭”状态的GDP，失去与标靶的亲和力而离开。

通常G蛋白在促进或刺激特异反应之前就识别靶点的特异构象。EF-G也如此，它能识别核糖体的易位状态，在GTP水解和Pi释放后在核糖体上解离。但是，EF-G是唯一与翻译有关且能与空的或停顿的核糖体相互作用的G蛋白，这也许是因为空的核糖体对易位前后的两个主要状态具有“构象记忆”[21]。这样，空的核糖体可以刺激依赖EF-G的GTP水解产生高反转，而EF-Tu却不能，这说明在一个低能源的最小化系统中，核糖体和EF-G都不能过量，当EF-G：核糖体的摩尔比为0.2～0.3时可达到这一条件，即每10个核糖体匹配2～3个EF-G分子。

2.7 结论

在这一章我们尝试估计了一个最小的活细胞所需的最小一套基因。对翻译装置最少组分的一个大胆估计是，其组成不超过200个基因，其中多于120个与翻译装置直接有关，约40个基因编码核糖体蛋白，两个rRNA（忽略5S rRNA）、21个tRNA、20个合成酶、6个因子和至少20个tRNA修饰酶。另外，最少30个基因在产生基本能量和合成某些氨基酸的过程中是必需的［另一些氨基酸在斯坦利·米勒模拟30亿年前大气和物理情况下可被产生[8]，它们可在最早的原生汤（primordial soup）中得到，而不需生物合成］。

即使我们只要一个“瘸腿的”生命形式，最后也需要至少150个基因。有趣的是，这和现在已知最低限度的生命形式，即木虱原生共生细菌一致，它有182个基因[22]。这种细菌生活在以韧皮部树浆为食的木虱体内的菌类细胞中，和大肠杆菌一样，属于γ-蛋白菌，通过宿主垂直传播而没有外源感染，它含有一整套翻译系统，并为宿主产生一些氨基酸。因此我们认为，使用远少于150个基因是不可能构建一个最小的细胞的。

参考文献

专题文献

1. Neidhardt FC. Chemical composition of Escherichia coli. In: Neidhardt FC et al, eds. Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology. Washington, D.C: American Society for Microbiology, 1987:3-6.
2. Andrade MA, OC, Sander C et al. Functional classes in the three domains of life. J Mol Evol 1999; 49:551-557.
3. Allwood AC et al. Stromatolite reef from the Early Archaean era of Australia. Nature 2006; 441:714-718.
4. Banerjee NR et al. Preservation of ~3.4-3.5 Ga microbial biomarkers in pillow lavas and hyaloclastites from the Barberton Greenstone Belt, South Africa. Earth Planet Sci Lett 2006; 241:707-722.
5. Eigen M et al. How old is the genetic code? Statistical geometry of tRNA provides an answer. Science 1989; 244:673-678.
6. Knoll AH. A new molecular window on early life. Science 1999; 285:1025-1026.
7. Gray MW, Burger G, Lang BF. The origin and early evolution of mitochondria. Genome Biol 2001; 2:reviews1018.1-5.
8. Miller SL. A production of amino acids under possible primitive Earth conditions. Science 1953; 117:528-529.
9. Alberts B et al. Chapter 14. Energy conversion: mitochondria and chloroplasts, in molecular biology of the cell, New York: Garland Science, 2002.
10. Chang BS, Halgamuge, Tang SL. Analysis of SD sequences in completed microbial genomes: Non-SD-led genes are as common as SD-led genes. Gene 2006; 373:90-99.
11. Moll I et al. Translation initiation with 70S ribosomes: An alternative pathway for leaderless mRNAs. Nucleic Acids Res 2004; 32:3354-3363.
12. Gualerzi CO et al. Initiation factors in the early events of mRNA translation in bacteria. Cold Spring Harb Symp Quant Biol 2001; 66:363-76.
13. Antoun A et al. How initiation factors maximize the accuracy of tRNA selection in initiation of bacterial protein synthesis. Mol Cell 2006; 23:183-193.
14. Dabbs ER. Mutant studies on the prokaryotic ribosome. Hardesty B and Kramer G, eds. Structure, Function and Genetics of Ribosomes. New York: Springer-Verlag, 1986:733-748.
15. O'Brien TW et al. The translation system of mammalian mitochondria. Biochem Biophys Acta 1990; 1050:174-178.
16. Umekage S, Ueda T. Spermidine inhibits transient and stable ribosome subunit dissociation. FEBS Lett 2006; 580:1222-1226.
17. Qin Y et al. The highly conserved LepA is a ribosomal elongation factor that back-translocates the ribosome. Cell 2006; 127:721-733.
18. Margus T, Remm M, Tenson T. Phylogenetic distribution of translational GTPases in bacteria. BMC Genomics 2007; 8:15.
19. Kim DM, Swartz JR. Prolonging cell-free protein synthesis by selective reagent additions. Biotechnol Prog 2000; 16:385-90.
20. Wilson DN, Nierhaus KH. The ribosome through the looking glass. Angew Chem Int Ed Engl 2003; 42:3464-3486.
21. Mesters JR et al. Synergism between the GTPase activities of EF-Tu.GTP and EF-G.GTP on empty ribosomes. Elongation factors as stimulators of the ribosomal oscillation between two conformations. J Mol Biol 1994; 242:644-654.
22. Nakabachi A, YA, Toh H et al. The 160-kilobase genome of the bacterial endosymbiont carsonella. Science 2006; 314:267.
23. Luria S, Gould S, Singer S. A view of life. Menlo Park, California: Benjamin/Cummings Pub. Co. 1981.
24. Dabbs ER. Mutants lacking individual ribosomal proteins as a tool to investigate ribosomal properties. Biochimie 1991; 73:639-645.
25. Alberts B et al. Chapter 2. Small molecules, energy and biosynthesis, in molecular biology of the cell. New York: Garland Science, 2002.
26. Madigan M, Martinko J, Parker J. Brock biology of microorganisms. 9th ed. London: Prentice Hall International, 2000.
27. Brocks JJ, LG, Buick R et al. Archean molecular fossils and the early rise of eukaryotes. Science 1999; 285:1033-1036.
28. Woese C. Interpreting the universal phylogenetic tree. Proc Nat Acad Sci USA 2000; 97:8392-8396.
29. Jewett MC, SJ. Mimicking the escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. Biotechnol and Bioengin 2004; 86:19-26.

3 行星上的太空生物学——外太阳系

François Raulin

3.1 行星太空生物学，可居住性及太阳系中的情况

太空生物学，是研究宇宙中生命的学科，更精确地说，是研究整个宇宙中生命起源、进化、分布及终结的学科。因此，这个非常广阔的跨学科领域不仅包括对宇宙生命的研究，也包括对我们这个星球上生命起源和进化的研究[1~3]。但我们必须记住，地球是迄今为止唯一已知有生命存在的星球，因而，显而易见的，地球上的生命是我们仅有的例子。为了寻找第二个例子，还需要探究的问题是，什么是生命，一个宇宙生命系统由何组成，以及如何探测到它[4]。

“什么是生命”是一个难题：如何通过我们所知的仅仅一个要素去定义一个整体？因此我们多数的研究都不得不基于这一个要素，将地球上的生命作为参考。通常认为生命是生活系统的特征，通过与环境交换物质和能量来保持高水平的信息量和复杂度，以实现自我增殖，适应外界压力，并通过自然选择而进化。在地球上，生命经过一段时期的化学进化而产生，通过复杂和大量的前生物化学进程实现，与液态水、有机物质和能量相关。结果产生的生活系统建立在碳化学之上，而它的发育和进化需要液态水、有机物质和无机营养。这些组分的可利用性似乎是生命起源和进化的必要条件。从地球上的例子开始，太空生物学将寻找地球外部以前或现在与地球相似的环境条件。

实际上，在太空生物学中，可以将我们最初对其产生兴趣的行星体进行分类。有些行星体上进行着复杂的有机化学反应。通常宇宙中物质复杂度的形成，在生命进化阶段是必需的，对这类化学反应进程和结构的研究，在理解这种过程时十分重要。在这个领域中，对彗星和陨石上有机化学的研究是首要的，因为它们的有机组成很可能直接构成了地球上前生物的化学组分。

还有一些行星体，与产生生命之前的我们的星球具有相似性。因为早期地球的多数条件在今天都已消失，被地质演变和生命自身抹去，所以对这种环境的研究有极大的重要性。现在，如果我们想了解地球上生命起源的进程，并验证我们的观点和概念，就必须将它们置于现实的环境中，现今可利用的行星体显示了与早期地球的可比性，这是一个难得的机会。在这个领域中，土卫六（Titan）——土星最大的卫星，是一个非常好的目标。

最后，还有一种太空行星体也许存在，其上生命或者已经消失，或者仍然存

在。这些行星体的特征是，在足够产生生命的时间内，具有能够发生复杂前生物进程的合适条件（或具有生活系统区别于其他系统的重要条件），并且具有适合其栖息的条件。其中一个决定行星体是否具有可居住性的重要参数是液态水的存在。火星，与地球类似，在其早期的很长一段时期，几亿年内，其表面具有大体积的液态水，这使这颗红色的星球成为在寻找地球外生命信号的痕迹方面，太阳系中最具吸引力的行星。实际上，如果生命曾经或仍然存在于火星上，现今也许可以在较近的地下表面上得到这些痕迹，因为火星的环境，虽然其大气经过了剧烈的变化，但由于没有强烈的构造更动，可能仍保留了某些生命痕迹。

但是，在太阳系中还有其他地方可能存在液态水。木星 4 颗“伽利略卫星”里的 3 颗——木卫三（Ganimede）、木卫四（Callisto）和木卫二（Europa）就存在这种可能（图 3.1）。同样地，土卫六（Titan）和最近才发现有这种迹象的另一颗较小的土星卫星 Enceladus 也有存在液态水的可能。尽管到现在为止我们还没有直接的证据证明内海的存在，不过木卫二（Europa）和土卫二（Enceladus）最能引起我们的兴趣，因为如果在这些星球上存在内海，那其内部水体可能和岩石材料结合在一起，促进氧化还原反应来产生保持前生物进程和生活系统的能量。接下来，我们将描述在太空生物学上最重要的外太阳系对象的状况[5,6]。

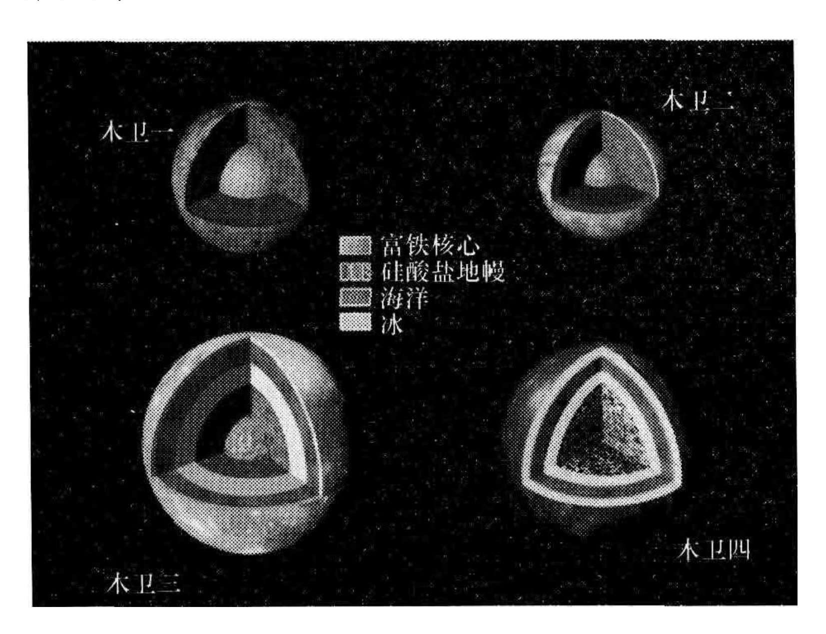


图 3.1　4 颗伽利略卫星的可能内部结构（根据 NASA/JPL，见彩插）。

3.2　类木行星系统：木卫二（Europa）

3.2.1　木卫二上的海洋

直径为 3 120 km 的木卫二是木星 4 颗最大的卫星之一，它于 1610 年被伽利略发现，也被称做“伽利略”卫星。在 Voyager 号在外太阳系的“大旅程”中，我们获得了大量关于大行星及其卫星的信息（图 3.2）。特别的是，根据 1979 年

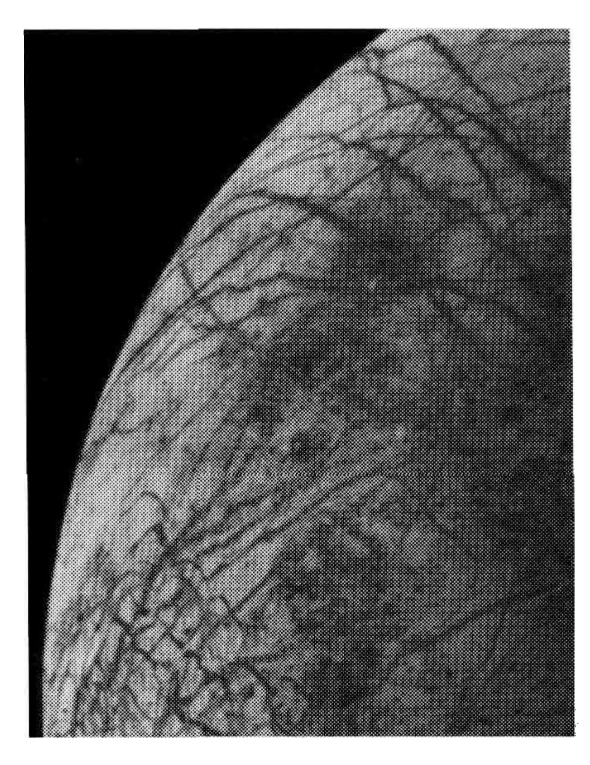


图 3.2 Voyager 2 距木卫二最接近时拍摄的图像，1979 年 7 月。其表面非常年轻，显示有复杂排列的条纹，这意味着它的外壳曾被内部释放的物质破裂并填充（来源：NASA/JPL）。

两艘 Voyager 太空船飞过木星时收集的信息，人们开始构建伽利略卫星内部结构的模型，部分人提出了它们中可能有内海，特别是在木卫二中（图 3.1）。不过绝大部分关于这些卫星的现有认识都来自于 1995～2003 年探索木星的伽利略任务。从木卫二卫星表面的高清晰照片中，我们能看到混乱的地表，很少的火山口，它们显示了地表重构的进程，还能看到许多很深的裂痕，显示了一个相对为液态的地下表面（图 3.3）。这些观测结果更加支持了内海存在的假设，并显示了内海可能离地表并不很远。木卫二的相对密度（相比于水）大概为 3，显示了它可能是由硅酸盐（相对密度 3～3.5）和水形成的冰（相对密度约 0.9）组成的。伽利略任务探测到了木卫二中的强感应磁场。这种现象可以有两种解释，那就是木卫二拥有一个铁核心，或者拥有超过 10km 深的水构成的一片海洋，并且它拥有和地球上海洋相似的传导性[7]。现有的可用数据还不足以让我们区分

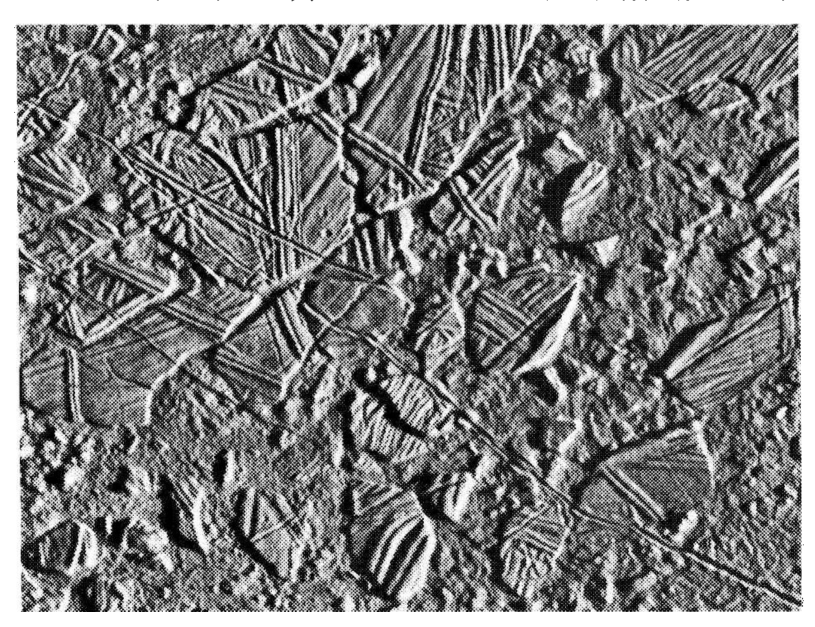


图 3.3 1997 年伽利略号拍摄的木卫二表面（约 34 km×42 km 的区域）的高分辨率图像，拍摄距离为 53 490 km。显示了地壳板块的分离和移动，这与地球极地冰的特征类似（来源：NASA/JPL）。

这两种可能，但是有一个被冰覆盖、(相对) 靠近地表的内海的可能性很大。

在过去的 30 年中，有许多有关这个项目的论文（见参考文献 7、8)。地表下的海洋可能有 100km 深，被大概 15km 厚的冰层覆盖，水温为 4℃[9]。最近，根据伽利略任务的数据，Hand 和 Chyba[10]重新检测了木卫二上海洋的可能性质。他们的一些主要结论显示，木卫二的海可能是接近饱和的硫酸镁水溶液，被厚度为 4km 或更薄的冰层覆盖。

由于其（不太深的）地表下很可能有液态水的存在，木卫二是一个引起太空生物学家极大兴趣的目标。在木卫二的早期历史中，很可能也存在生命产生所必需的其他条件（可利用的有机物质和能量)。与地球上相似的是，陨星上的许多有机质也许对木卫二非常重要。此外，伽利略卫星也许在很大程度上是由含碳陨星上的物质构成，如CⅡ球粒状陨石。它们含约 2.5%的碳和 13.5%的水。由于木卫二的含水量可能占其质量的 7%以上，如果它由 CⅡ球粒状陨石构成的假设是正确的，那么它可能也含有占其质量 1%的碳原子[11]，其中一半为不可溶的复杂的有机物。后者可能在经历非均匀的水解过程后，产生与生命相关的有机物。水溶性的有机物碎片可能约占木卫二质量的 0.05%，并且包括许多与生命相关的组分。

假设所有的水都呈液态，则可以估计出木卫二的内海中，溶解的有机碳浓度可高达 1%。这样的有机物浓度，由于其性质，可能足以产生高效的前生物进化。举例来说，水溶液中浓度为 0.1mol/L 的 HCN，足以聚合产生 HCN 的四聚体及更高级的寡聚体，通过水解，成为氨基酸及嘌呤和嘧啶碱基的来源。低温降低了前生物化学反应的一般速率，但也可能通过共熔效应增加反应物浓度，从而提高反应速率。另外，如地下海这样的高压条件也可能会促进化学液化反应，而这些反应对生物大分子通过其单体进行构成是必需的，如多肽和多核苷酸分别由它们的结构单元氨基酸和核苷酸构成。

此外，在这些海洋的底部可能具有热液口，为增加化学复杂性提供了有利位置，这得益于能在高温气体及液相和固相间界面上发生的异质过程。这种过程可以被无机相潜在的催化特性促进，同时也被很高的热梯度促进，它可以保护产物不被热降解。但是，最近 Hand 和 Chyba[10]对海洋盐度的研究显示，高浓度的盐分可以在相当程度上降低低聚反应，这种反应对前生物出现、大分子的复制是必需的。这样看来，这些海的现有条件对生命起源不是非常有利，但的确不能排除后者可能是在海洋条件变得不同之前发生的。

3.2.2 木卫二的可居住性

如果生命已经在木卫二上的海中出现，人们有理由质疑这种环境的可居住性。首先，邻近木星上的高能辐射可能是生命存在的主要问题。但数千米厚的冰壳是一个很好的保护罩。海洋的温度（似乎在 0℃附近或更高)[9] 不应是对生命发

展的问题。如果我们考虑到地球上更极端环境中嗜极生命体的存在，那么潜在的高盐浓度，可能的大范围 pH 和高压（最大为 100 MPa）并不是障碍。因此不能排除在木卫二上的海洋中有生命系统的存在[11]。

木卫二上假定的生物群有多广泛呢？Reynolds 等人[12]检测了木卫二上生物系统可利用的不同的可能能源。考虑到只有太阳能是可用的，他们估计可能的生物量为 $2\times10^{-6}g/m^2$。这个值与地球上较低水平的生物量（1000～10 000 g/m^2）相比也是相当小的。虽然如此，尽管非常有限，它也是一个不可忽略的生物群，此外，这种估计仅仅是基于太阳能的，因此，如果考虑到这颗伽利略卫星上其他能源的存在，木卫二上现有的可能生物量也许更高，尤其重要的是由放射能产生的热流传递。实际上，考虑到在冰和海洋中放射性的 ^{40}K 的衰变及 1000 年间生物量的流动而额外产生的氢和氧，Chyba 和 Phillips[8] 估计木卫二上稳定的生物量水平为 $10^{13}\sim10^{15}g$。在这个行星体表面，给出的平均生物量密度为 $0.3\sim30g/m^2$。

我们可以期望出现什么样的生命呢？我们不能排除真核生物的进化，但原核厌氧的生命更为可能。类似古菌的有机体可能是木卫二上最有可能存在的生物群[13]。地球上的古菌实际上是能够预期的最好例子。表 3.1 显示了大量具有不同代谢类型的可能的例子。

表 3.1　可能对木卫二有意义的地球生物体的例子（根据 Oro 等人的相关研究，见参考文献 11）

产甲烷古细菌
异养（发酵）：有机物＋H_2＝>CH_4＋生物合成产物
有机物＝甲醇，甲胺，甲酸盐，乙酸，…
自养（CO_2还原）：CO_2＋H_2或 Fe^0＝>CH_4＋生物合成产物
嗜热古细菌
异氧：（发酵）：酵母提取物＝>CO_2＋生物合成产物
（S 呼吸作用）：有机物＋S＝>H_2S＋CO_2＋生物合成产物
自养（S 还原）：CO_2＋H_2＋S＝>H_2S＋生物合成产物
有机物＝乙醇，糖，甲酸盐，乙酸，…
光合细菌
光合异养（厌氧）：有机物＋光＝>生物合成产物
光合自养（厌氧）：CO_2＋$2H_2S$＋光＝>S_2＋H_2O＋（HCHO）
光合自养（有氧）：CO_2＋H_2O＋光＝>O_2＋（HCHO）

3.2.3　对木卫二的太空生物学探索

我们怎样才能探测到木卫二上的生命？一个强有力的手段是“水栖机器人/冰中机器人”任务，用一个冰中机器人的熔化器穿过数千米厚的冰壳，释放一个

潜水器（仍是假设）进入木卫二的海洋（图 3.4）。这个潜水器可以装备完善，以寻找可能的热液口及其附近可能存在的微生物。但是，有许多与这个任务相关的问题，尤其是资金和许多技术上的难题。特别是它需要将地球上的技术在太空中应用，这种适应非常困难。地球上南极洲 Vostok 湖的例子对这个问题给出了一个清晰的例证。经过多年的钻探，几乎能够到达位于在 3km 厚的冰层下的这个地下湖的顶端。但是，由于生物污染这个严重的问题，科学家们现在在犹豫是否继续深入。由于更低的生物量密度，这个问题在地球外的液态水体，如木卫二的海洋中，将更严重。实际上，在假设的木卫二海洋中，生物活性可能十分有限并难以探测，而潜在的生物污染的重要性更高一些。



图 3.4 用于探测木卫二上的海的冰下机器人示意图（来源：NASA/JPL，见彩插）。

因此，在当前的情况下一个在轨道运行的任务似乎更为合理。它可能通过 NASA-JIMO（Jupiter Icy Moons，木星冰月亮）任务得以实现，这个任务计划提供一个木卫二的详细地图，并提供数据来证实内海的存在，这个任务最近由于资金短缺而被取消。其他关于木卫二的任务正在 ESA 宇宙宏观给出的可能性框架中进行研究，这些任务将能够证明一个木卫二内海的存在，并且在这颗卫星表面，通过寻找代谢活动产生的分子信号，如 CH_4、H_2S、HCHO 等，来探寻生命活动的印迹。

3.3 土星系统：土卫六（Titan）和土卫二（Enceladus）的情况

3.3.1 一些历史背景和一般数据[6]

在 1655 年由荷兰天文学家 Christiaan Huygens 发现后，这颗土星最大的卫

星在很久之后的 19 世纪被 John Herschel 命名为“土卫六”（巨人）。环绕土卫六的可视圆面，有一个发暗的边缘。1908 年，西班牙天文学家 Jose Comas-Sola 对其研究显示，土卫六周围存在可以探知的大气。英国天文学家 James Jeans 仔细考虑了这个假说，并计算出，若温度足够低（低于 100 K)，一个类似土卫六的行星体可以维持一个实体的大气（多数主要成分的分子质量都超过了 16D)，而不会因为它相对较小的体积（直径 5150 km）使其逃逸。1944 年，通过美国天文学家 Gerard Kuiper 对气体甲烷（CH_4）吸收带的检测，明确证实了土卫六周围大气的存在。他接着论证了甲烷为其大气主要组分之一，甚至还成功地由他的观察推论出 Titan 表面甲烷的分压应为 0.1 bar（100 hPa）左右。之后，通过许多地面望远镜和一些经过土星的太空任务的观察，证实了对 Titan 大气中甲烷的检测结果。

半径为 2575 km 的土卫六（图 3.5）是木卫三（R＝2631 km）之后太阳系中第二大的卫星。以土星来说，它环绕太阳的运行周期约为 30 年，到太阳的平均距离约为 9.5AU，其对应的太阳辐射流水平约是地球轨道水平的 1%。如果将土卫六看做一个黑体，它的平均温度为 82K。实际上，土卫六是太阳系中唯一具有可观察到的大气的卫星，同时也是唯一与地球具有相似大气的行星体。诚然，土卫六的大气主要由氮气（N_2）组成，且其表面压力为 1.5 bar（1500 hPa）。如前文所述，它的大气也含有一定量的甲烷。土卫六的表面温度约为 94K，由于显著的温室效应（20K，主要由大气中的甲烷和二氢化物产生)，虽然有由雾粒子

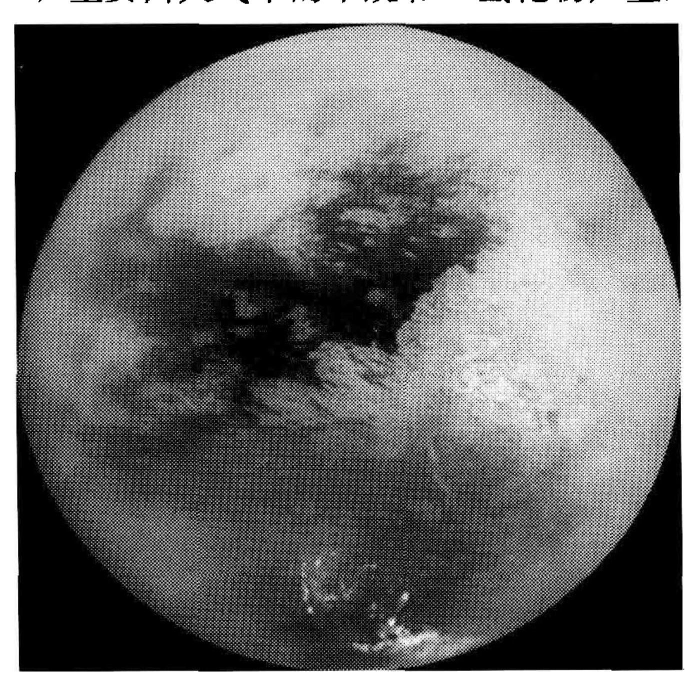


图 3.5　土卫六表面的镶嵌图案，由卡西尼飞船上的窄角相机拍摄。大块的亮白区域被命名为 Xanadu。分辨率约 1.3 km/pixel（来源：NASA/JPL/Space Science Institute)。

产生的逆温室效应（约－8K），但仍比黑体温度高了约 12K。事实上，由于表面温度如此之低，土卫六表面附近的大气密度比地球大气高约五倍。

土卫六以 16 个地球日为周期绕土星旋转，同时自转，因此，土卫六的固体表面缓慢自转，但它的大气由于强大的带状气流而呈超回转现象。由于到土星的距离约等于 20 个土星半径，土卫六距这颗大行星足够远，使其能够避免与土星环作用；也足够近，使它的大气能与土星磁性层的离子进行相互作用，它们在化学演化中起重要作用。

3.3.2 理论模型

土卫六的轨道有轻微的偏心（e=0.029）：这意味着土卫六的表面不存在浅海，因为它将减小偏心值，使轨道变为圆形。现在，土卫六的平均密度为 $1.88g/cm^3$。这么低的值，说明它是由低密度物质——冰（如平均密度约为$1g/cm^3$的水冰）与高密度物质（如平均密度为 $3g/cm^3$的岩石和硅酸盐）混合组成。

约 46 亿年前，土卫六起源于土星次影中一个小行星体的增长。由重力过程和放射性衰变产生的热量融化了冰组分，使岩石向形成中的土卫六的中心陷落。结果最初的原始结构是，一个由硅酸盐、岩石和冰充分混合构成的中心核，被硅酸盐外核和液态的氨水混合物覆盖，再与一个高密度的大气接触（图 3.6）。关于这个原始大气的起源和组成仍有争论。最近关于土星次影演化的模型显示，形成土卫六的小行星体包括 CH_4 和 NH_3 的水合物形式（不是 CO 或 N_2，它们不能有效形成水合物）。接着 CH_4 和 NH_3 逐渐被释放到大气中。NH_3 在大气上层区域经光化学过程转化为 N_2，或经电击作用催化分解。那时土卫六的表面足够

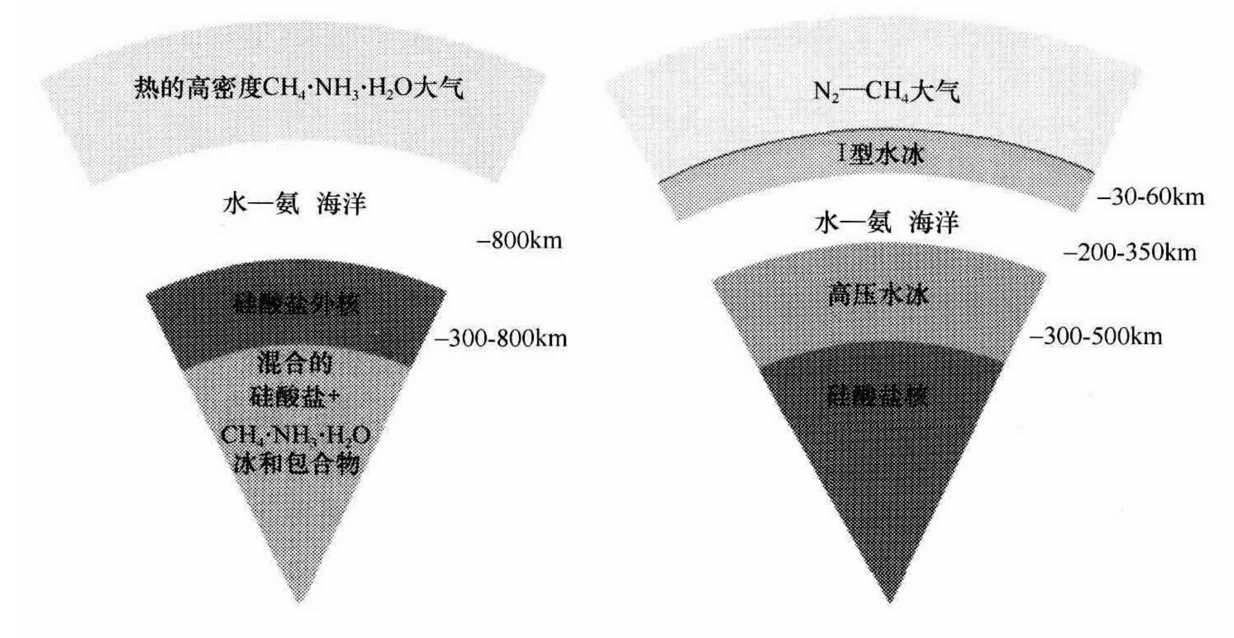


图 3.6 土卫六的内部结构；左：原始土卫六；右：现在的土卫六（根据 Icarus, Vol. 146, Fortes[24]；2000 年版权，得到 Elsevier 的许可）。

温暖，可以维持水环境达数千万年（可达 1 亿年）。随着土卫六的能量减少，它逐渐冷却，液态水开始由 H_2O-NH_3 溶液结冰。一个Ⅰ型冰壳（Ⅰ型冰为地球上所有自然形成的雪和冰的形式）形成，它的厚度在 7 千万年里迅速增长到 30 km，覆盖着一个最初与底层岩石相接触的内海。底部的高压导致了Ⅵ型水冰（高压水冰）和氨水合物的结晶化，使内海富含氨水。海的深度和组成取决于流动和对流的热能；它可能有 200 km 厚，含大于 15 wt%的 NH_3。

根据大气的演化和当前情况，已发布了几个动力学模型，主要是光化学模型。它们主要给出了大量及微量物质种类的垂直浓度分布。由于在卡西尼-惠更斯任务之前缺乏对涡流扩散系数的详细观察和数据，这些分布数据相对来说非常局限。但是，对与这些复杂的大气化学作用相关的化学进程，这些模型提供了初步解释。它们假定，这些化学过程开始于太阳紫外线光子及土星磁性层的电子对 CH_4 和 N_2 的降解。生成的初级物质在高层大气中形成，产生简单的碳氢化合物和含氮分子，特别是 C_2H_2 和 HCN。一旦形成，这两种组分就下降到底层大气区域，在那里吸收可利用的紫外光子（在近紫外区），因为它们的紫外光谱可以扩展到那个区域。尽管能量低于中紫外光，但这些光子也可以导致发生光解离。生成的初产物产生更复杂的碳氢化合物和腈。在化学反应过程中，聚炔烃（如 C_4H_2 和 C_6H_2）及多炔氰胺（如 HC_5N）可能在形成高分子质量的产物中起重要作用。其他的 CH_4 降解可能在低同温层中，通过与 C_2H_2 和聚炔烃相关的光催化反应进行。这个化学反应的终产物将是一个主要组分为 C、H 和 N 的高分子有机物。然而，如下所述，通过卡西尼-惠更斯任务[15,16]对土卫六的电离层进行的改进后的测量显示，在高层大气中进行的有机过程一定也参与了这个高分子物质的形成。但是，与这个光化学模型偶联的详细的离子化学模型仍需制订。

3.3.3　实验室实验

另一个研究这些进程的手段为实验室实验。为得到一些与行星体进化相关的参数，如速率常数、热力学数据或光谱数据，做一些特别设计的实验。也有一些实验部分或整体模拟特定环境的演化。所谓的“模拟实验”可以综合许多不同的相关理化过程。其中最早在实验室中进行的模拟实验，是历史上斯坦利·米勒模拟早期地球环境化学演化过程的实验。由于缺乏任何对该环境的直接观察数据，该实验事实上是研究的唯一工具。几组研究人员近来较多使用了类似的手段，最初的反应装置是密闭的，最近则使用开放的流动反应装置，使用等离子体放电作为能源，并用分子氮和甲烷的流动混合物，来研究土卫六大气的演化和有机化学过程（图 3.7）。

在土卫六的大气中，主要的能量来源是太阳紫外线辐射和来自土星磁性层的

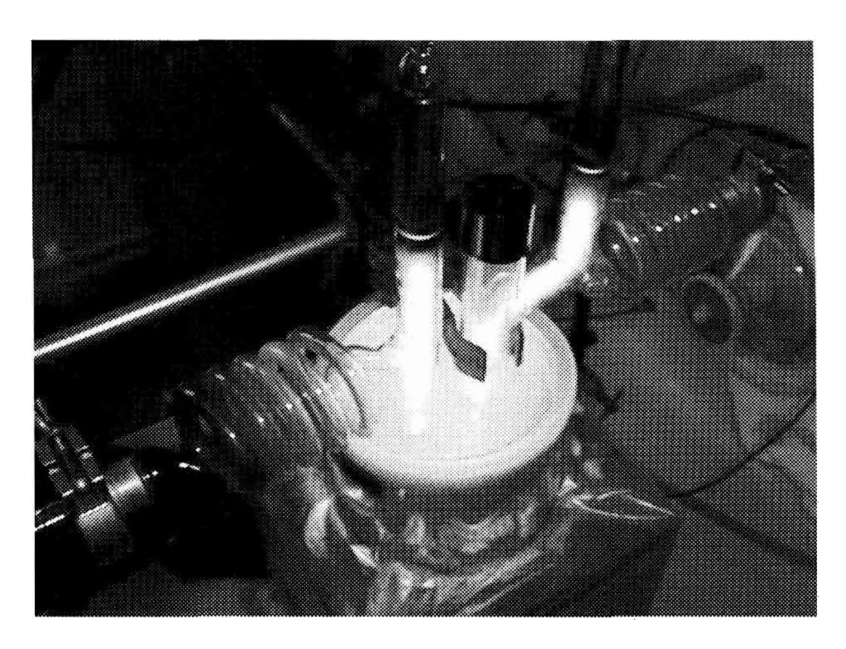


图 3.7 LISA (Laboratoire Interuniversity des Systèmes Atmosphériques，巴黎大学 12 及 7) 对土卫六大气化学反应进行实验室模拟的试验装置。N_2 (98%) 和 CH_4 (2%) 的混合气体连续流经一个开放的反应器，并维持低压 (约 1 hPa) 和低温 (约 150 K) (来源：P. Coll)。

中能电子。在中紫外光中最丰富的太阳光子，是 Lyman α 光子 (121.6 nm)，与甲烷光解离 ($\lambda < \sim 150$ nm) 相一致。在实验室中可以用单色光轻易模拟这种辐射，但 Lyman α 光子不引起氮气光解离，它需要波长更短的光子 ($\lambda < \sim 100$nm)。用与密闭或流动反应装置相一致的发光系统，在实验室中难以得到这种远紫外光。因此，多数已经进行的，用 N_2—CH_4 为起始气体混合物，模拟土卫六大气化学演化的实验室实验，都使用电子碰撞能源，以模拟来自土星磁性层的电子对土卫六大气的撞击。另一个方法与短波紫外线能源相一致，利用含氮组分的初始气体混合物，它可以被波长较长的紫外线光子光解离，且也存在于土卫六大气中，尽管是在微量水平，如 HCN 和 HC_3N 分子。

自 20 世纪 80 年代早期以来，许多实验室实验都模拟过土卫六大气的化学过程，已在气相检测出超过 200 个不同的有机分子，主要是碳氢化合物和腈。在这些实验中，所有已在土卫六大气中观测到的气体有机物种类，与其他许多 (仍) 未观测到的，都被检测了出来，尤其是像 C_6H_2 和 C_8H_2 这样的聚炔烃，以及 HC_4—CN 这样的多炔氰胺。这些组分在土卫六大气的光化学模型中也存在，可能在气相产物向浮质转化的化学过程中起关键作用。之后的太空生物学研究方向为不均匀的含碳有机组分及当 CO 存在时 (在土卫六大气中以 100 ppm 水平存在) 含氧组分的形成，其中含量最多的含氧有机物为环氧乙烷 (又称乙烯化氧) $(CH_2)_2O$。

模拟实验也产生固体有机物，通常称为多林 (tholin)，这是一个在 20 世纪

70 年代由 Carl Sagan 命名的通用性名称[17]。这些“土卫六多林”应该是实验室中土卫六浮质的类似物。自从 20 年前 Sagan 和 Khare 进行最初的研究以来，它们已经被广泛研究。这些实验室中的类似物根据实验条件显示出非常不同的特性。举例来说，产物的平均 C/N 在数值上可能从小于 1 到大于 11。仍然几乎不知道土卫六多林的化学组分，但可以确定它们由结构上非常不同寻常的高分子组成。红外、紫外光谱和热解-GC-MS 技术分析显示了脂肪族、芳香族碳氢化合物及 CN、NH_2 和 C=NH 基团的存在。它们的水解，即使在中性 pH 条件下，也释放出氨基酸；同样具有太空生物学重要性的是，一些微生物（地球上的）可以利用它们作为氮源。

3.3.4　观察：卡西尼-惠更斯任务的情况

诚然，上述两种研究方法，即建模和模拟实验，没有一种可以在没有任何观测数据的情况下发展，如 NASA-ESA 卡西尼-惠更斯任务提供的数据，它结合了遥感技术和就地观察[15,16]。卡西尼太空船，携带着惠更斯土卫六大气探测器，于 1997 年 10 月 15 日发射，在行星间轨道上运行 7 年后，于 2004 年到达土星，在 2004 年 7 月 1 日并入土星轨道，使卡西尼号成为土星的一颗新的卫星。在 2004 年末，在与土卫六相遇两次之后，它于 2004 年 12 月 25 日将惠更斯探测器释放到了环绕土星的第三个轨道上。在 2005 年 2 月 14 日，惠更斯探测器穿过了土卫六的大气。经过了 2.5h 的降落，惠更斯探测器着陆，并在其表面上继续运行了 3h。在离开着陆的探测器的视野之前，卡西尼号收集行星表面数据达 1 小时 10 分钟以上。

对土卫六的探测是卡西尼-惠更斯任务最主要的目的之一。名义上 4 年的任务，包括了 74 个环绕土星的轨道器，其中 44 个是土卫六的近天体探测卫星。卫星 12 个探测仪器中的一部分（表 3.2A）及表面探测器 6 个探测仪器中的多数（表 3.2B)，都提供了太空生物学相关的数据。在卡西尼号上，离子和中子质量分光仪（INMS)、复合红外分光计（CIRS）和紫外成像摄谱仪（UVIS)，确定了土卫六大气不同区域的化学组分，特别是它们能探测到许多包括一些新物质种类在内的有机物，并且能确定它们的垂直浓度分布。卡西尼雷达、科学成像系统（ISS）上的摄影机，以及可见和红外绘图光谱仪（VIMS）可以通过霾层测绘土卫六的表面，提供有关其表面形态学、地质学和化学组分的信息。

在惠更斯探测器[15,18]降落中及着陆后[19]，探测器上的 GC-MS 仪完成了一个对其大气的详细的化学分析。在探测器降落过程中，浮质收集和热解（ACP）实验设法收集到了土卫六上的浮质，并将它们加热到不同程度的高温，以对收集到的粒子中耐高温的部分进行热解。生成的气体被送到 GC-MS 仪中进行分子分析，这是第一次直接对土卫六浮质进行的原位分子和元素分析[20]。特别是惠更斯

表 3.2A 卡西尼号上的仪器,跨学科研究项目(IDP),首席科学家和获得太空生物学发现的可能性

设备或跨学科项目	P. L. ,T. L, 或 IDS'	国家	太空生物学 返回情况
光遥感设备			
复合红外分光计(CIRS)	V. Kunde, M. Flasar	USA	+++
科学成像系统	C. Porco	USA	+++
紫外成像摄谱仪	L. Espsito	USA	++
红外绘图光谱仪	R. Brown	USA	++
场粒子和波设备			
卡西尼等离子光谱仪	D. Young	USA	+
宇宙尘埃分析	E. Grün	Germany	+
离子和中子质量分光仪	H. Waite	USA	+++
磁力计	D. Southwood, M. Doughert	USA	NA#
磁层成像仪	S. Krimigis	USA	NA
无线及等离子波谱仪	D. Gurnett	USA	NA
微波遥感			
卡西尼雷达	C. Elachi	USA	+++
无线科学系统	A. Kliore	USA	++
跨学科项目			
磁层和等离子体	M. Blanc	France	+
环及尘埃	J. N. Cuzzi	USA	+
磁层和等离子体	T. L. Gombosi	USA	+
大气	T. Owen	USA	+++
卫星及小行星	L. A. Soderbolom	USA	+
高层大气与太阳风的作用	D. F. Strobel	USA	++

注:P. L. =主要负责人 T. L. =队长 IDS=跨学科科学家 #NA=不适用

表 3.2B 惠更斯探针上的仪器,跨学科研究项目(IDP),首席科学家和获得太空生物学发现的可能性

设备或跨学科项目	P. L. ,T. L, 或 IDS'	国家	太空生物学 返回情况
气相质谱	H. Niemann	USA	+++
浮质收集和热解	G. Israël	France	+++
惠更斯大气结构仪	M. Fulchignoni	Italy	++
降落成像光谱辐射仪	M. Tomasko	USA	+++
多普勒风仪	M. Biro	Germany	+
表面科学工具包	J. Zarnecki	UK	+++
跨学科项目			
高层大气物理学	D. Gautier	France	++
大气/表面相互作用	J. l. Lunine	USA	++
化学及太空生物学	F. Raulin	France	+++

注:P. L. =主要负责人 IDS=跨学科科学家

大气结构仪（HASI）确定了压力和温度的垂直分布。降落成像光谱辐射仪（DISR）测量了大气的辐射收支，研究了云层结构并拍摄了行星表面图像（图3.8)。表面科学工具包（SSP）提供了行星表面物理情况和化学组分的信息。通过对潜在的时空变化的观察，来自卡西尼卫星的数据给出了土卫六的全局图谱。惠更斯探测器的数据对土卫六的一个特定区域给出了非常详细的信息。虽然卡西尼任务还远远没有完成，由轨道卫星和降落探测器提供的互补合作已经显示了对土卫六的新的视野，在太空生物学主要体现为三个方面：① 与地球，包括早期地球具有许多相似点；② 复杂的有机化学组成；③ 潜在的可居住性。

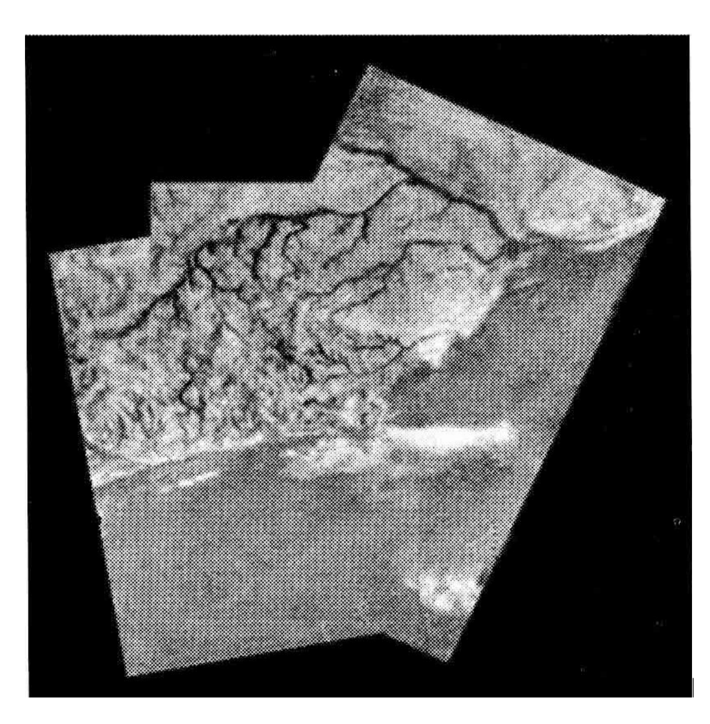


图 3.8 土卫六表面的照片，由惠更斯探测器上的DISR仪在6.5 km高处拍摄，显示了河道网络、高地和明暗分界面（来源：ESA/NASA/JPL/亚利桑那州立大学）。

3.3.5 土卫六与太空生物学相关的主要方面

土卫六的高密度大气，延伸到约1500 km，主要由N_2组成，与地球大气相似。其他的主要组分为CH_4（1.6%～2%，同温层，由卡西尼上的CIRS和惠更斯上的GC-MS仪测量得到）和H_2（约0.1%)。虽然土卫六更冷一些，对流层（～94—～70K)，对流层顶（～70.4K)，同温层（～70—175K)，最近经卡西尼-惠更斯任务证明，它的大气呈现与地球类似的复杂结构，并且也包括一个中间层和一个热对流层。这些相似性与这两个大气中温室气体及逆温室气体的存在相关。在土卫六上，CH_4和H_2与地球上可冷凝的H_2O及不可冷凝的CO_2等价。另外，土卫六大气中的薄雾粒子和云与地球大气中的浮质和云类似，起逆温室效

应的作用。土卫六上的甲烷看起来与地球上水的角色类似，其复杂的循环还没有完全阐明。卡西尼 ISS 探测了南极附近的地表特征，那里可能有一个湖；雷达已经在北极区域发现了湖和海的存在，可能是液态甲烷和乙烷（图 3.9）。而且，惠更斯上的 DISR 仪提供了土卫六表面的照片，非常清楚地显示了类似地球河流网络的树状结构（图 3.8）。没有陨石坑，这些特征体现了一个液体开始流动不久的相对年轻的地带。由惠更斯 GC-MS 提供的数据显示，在着陆点上空约8 km 高处的对流层中，CH_4 含量达到饱和水平，具有形成云的可能性。结果可能带来的雨也许是观测到的河流状网络的成因。此外，GC-MS 数据也显示了在着陆点存在有可冷凝的甲烷。其他卡西尼号上仪器的观测清楚地显示了存在各种不同表面特征，表明火山、构造学、沉积学、气象学过程的不同起源，它们与地球上的这些过程类似（图 3.10）。

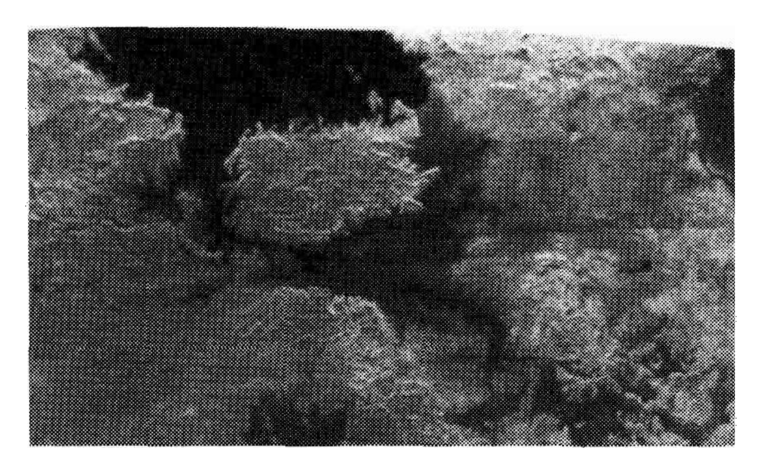


图 3.9 土卫六表面的照片，于 2007 年 2 月由卡西尼雷达在高纬度地区获得，显示了一个大湖的存在，约 100km，有一个约 100km 大小的岛（来源：NASA/JPL）。



图 3.10 上部：卡西尼在土卫六表面观察到的沙丘（上部）看起来与地球上的 Namibian 沙丘（下部）类似（来源：上方照片 NASA/JPL；下方照片 NASA/JSC）。

INMS和GC-MS已探测到土卫六大气中氩气（Ar）的存在。类似于地球大气，丰度最高的同位素是^{40}Ar，应该来源于放射性衰变的^{40}K。由GC-MS测定，它在同温层中的摩尔分数约为4×10^{-5}。原始的Ar（^{36}Ar）含量约比它小200倍。其他原始稀有气体的混合比小于10 ppb（$1ppb=10^{-3}ppm=10^{-9}$），这在很大程度上意味着土卫六大气来自第二级起源，从固体中被捕获的气体释放出来。而且，N_2是不可以被这颗形成土卫六的冰冻小行星体有效地捕获的，但NH_3可以。这样，人们认为土卫六的原始大气最初由NH_3组成，后来通过光解和（或）撞击引起的化学过程转化为N_2。同温层中，由GC-MS测量的$^{14}N/^{15}N$比例（=183），比原始的氮中小1.5倍。这说明，自形成以来，在甲烷转化成有机物的同时，大气可能有几次散失，意味着土卫六表面存在大量有机物沉积。

3.3.6 活跃的有机化学反应

我们也可以将土卫六上正在进行的活跃的有机化学反应，与早期地球上的前生物化学进程进行类比。在土卫六的大气中已经探测到了一些有机物，它们在地球前生物化学过程中通常被认为是关键分子，包括氢氰酸（HCN）、丙炔腈（HC_3N）和氰（C_2N_2）。同温层中这些被探测到的有机物[21]为碳氢化合物（含饱和和不饱和链）和含氮有机组分，主要是腈，这与实验室模拟实验预期的一致。自从卡西尼号到达土星系统，CIRS仪已经明确证实了水和苯的存在。

在卡西尼号低空飞过土卫六时，INMS对电离层的直接分析显示，在非常高的高度（1100～1300km）许多有机物都以可探测的水平存在。这个仪器可探测的质量范围最大值是100 Da。但是，对数据的推测在很大程度上显示，质量达几千道尔顿的高分子物质很可能在电离层中存在[22]，这些新的数据打开了对土卫六大气中正在进行的有机化学过程的全新视野，强烈暗示着在土卫六环境中电离层化学反应对复杂有机物形成的作用，而在此之前没有想到过这一点（图3.11）。

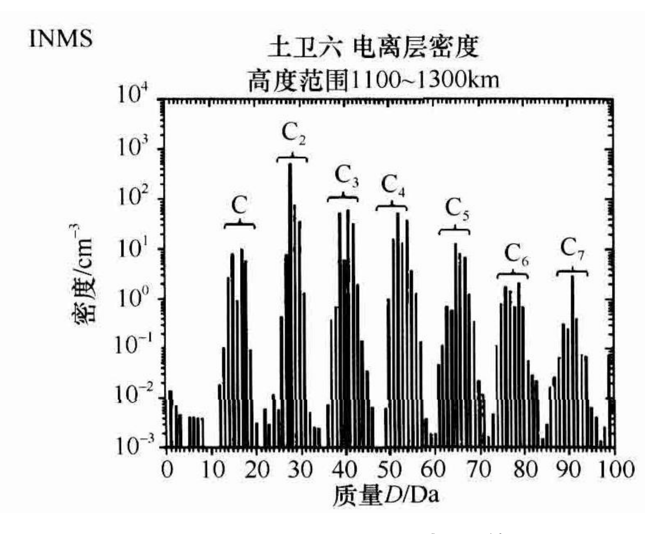


图3.11　土卫六电离层的平均质谱，由卡西尼-INMS在1200 km高处得到。质谱显示了包括多达7个（碳/氮）原子的有机化合物的特征（来源：NASA/JPL/密歇根州立大学）。

与此相反的是，惠更斯上的GC-MS仪[19]在低层大气中没有探测到许多有机物。这有可能是因为它们都冷凝在浮质中。由于以前没有关于它们化学组成的直接数据，这些粒子由ACP仪进行分析[20]。数据显示，这些浮质由耐高温的

有机物组成，在热解时释放出 HCN 和 NH_3。另外，这些热解物显示可能存在腈、氨基酸和（或）亚氨基基团。这个结果有力支持了“浮质的分子组成和实验室多林非常类似”这个假说。这些粒子的可能组成为一个耐热的有机核被挥发性的可冷凝组分包被，平均直径为 1μm（图 3.12）。

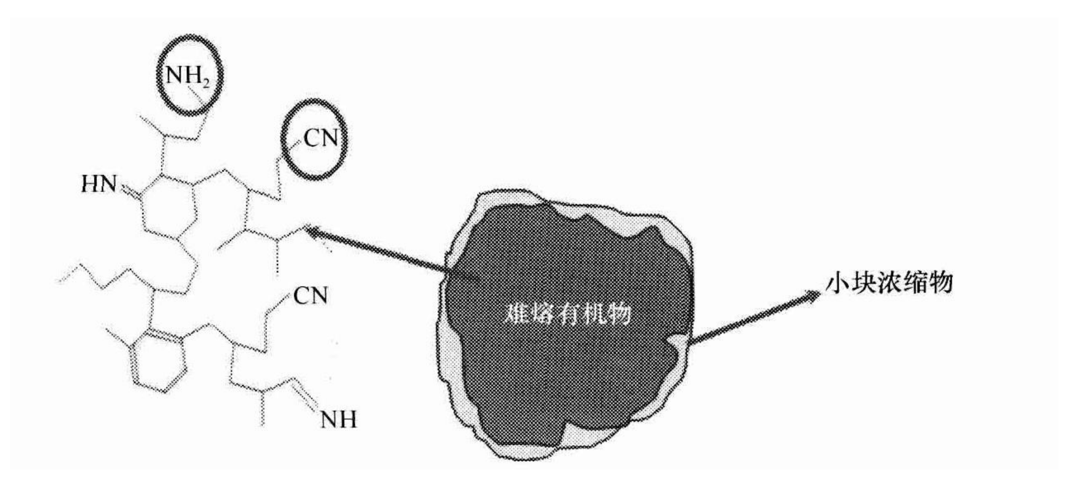


图 3.12 左图：土卫六浮质的化学组分模型（由惠更斯 ACP 数据推测得到）。

在沉积后，浮质应积累在行星表面，形成复杂的耐热有机物和冰冻的挥发性物质的堆积。DISR 数据显示了多林中水冰的存在，但尚且没有确切的证据。此外，在惠更斯号着陆后，GC-MS 对地表附近大气的分析显示，可能存在多种有机物，包括含氮有机物，C3、C4 碳氢化合物和苯，它们被探测器上的热量蒸发了出来。这与以下假说一致，即星体表面富含可冷凝的挥发性有机物，且这些有机物中的多数都以冷凝相存在于低层大气中。

基于对土卫六中有机化学反应的这些了解，大气中主要成分，氮气和甲烷的化学演化主要产生乙烷，它积累在地表或近地下处，最终溶解在甲烷-乙烷形成的湖和海中，同时溶解的还有积累在地表的复杂耐热有机物及冷凝的挥发性有机组分，如 HCN 和苯。虽然温度很低，与通常概念相反的是，土卫六并不是一个冻结的地球，其化学系统并没有冻住。土卫六是一个进化中的行星体，它的化学反应也在进化中。一旦沉积在土卫六表面，浮质和其中的复杂有机成分可能接下来就会进行与太空生物学相关的化学演化。实验室实验表明，土卫六多林和液态水接触可以释放许多与生命相关的组分，如氨基酸，和水冰接触亦可以。这些过程尤其可能在土卫六表面有冰火山的区域进行。这样我们可以预想到这些组分可能存在于土卫六表面或近地下处。因此对土卫六表面的就地测量可以提供一个唯一的机会，通过地面实况调查，研究许多可能与前生物化学反应相关的过程，包括那些同位素和对映体片段。

实际上，尽管土卫六表面温度非常低，液态水仍有可能在其表面存在。彗星

冲击可以融化其表面的水冰，并在长达 1000 年的时间里，通过局部的液态水体，为与地球类似的前生物合成提供条件[23]。前面假设存在的内部水-氨海洋也是一个可以发生前生物化学反应的地点，这与木卫二的情况类似。由于土卫六的海洋可能不再与硅酸盐层接触，这些前生物进程发生的可能性减小，但仍可能在土卫六的早期历史中发生，这是因为可能存在的高温液体活性，使得 C、H、O、N（碳、氢、氧、氮）前生物化学反应可以在很长时间（几百万年）内进行，因此，还不能排除现在在这片隐藏的海洋中可能有生命存在。

3.3.7 土卫六的可居住性

实际上，虽然土卫六的表面温度过低，这个环境里可利用的能源不足以提供可居住的条件，但它的地下海可能适合生命生存。如 Fortes 所说，海的可能平均温度（260K，在冰火山的热点附近甚至可达 300K）并不是生命系统发展的障碍。这同样适用于压力和 pH 参数。在 200 km 深处，预计土卫六海洋的压力为 500 MPa（5 kbar），并不像地球海洋那样不适合生命。其 pH（水溶液中 NH_3 质量可达～15%）约为 11.5，在地球上，细菌可以在 pH 为 12 时生长。我们同样不能排除有限的能源不可以维持生命的需要。

据 Fortes 估计，土卫六的海洋中可能约有 5×10^8 W 的能流可供生命活动使用。在地球上的生物圈中，与生产相对应的这个值每年约为 4×10^{11} mol 的 ATP 和约 2×10^{13} g 的生物量。假设平均每年生物系统以这样的速度流通，生物量密度将达到 1g/cm，这几乎达到了木卫二上海洋的水平。卡西尼-惠更斯数据不能证实土卫六表面有任何生命活动的迹象。尤其是，对甲烷碳同位素比例的测量明显表示并没有生命产生。但是，这些数据不能排除在土卫六地下海洋中可能有有限的生物群存在。

3.3.8 土卫二（Enceladus）的惊喜

当然，卡西尼-惠更斯任务尚未完成，但仍可以给出令人惊奇的发现，如土卫二。在卡西尼-惠更斯任务之前，对这颗土星的小卫星（半径只有 250 km）所知甚少。自从卡西尼卫星对土卫二开始观测以来，这颗令人惊奇的卫星正在成为太空生物学行星研究的新目标。

卡西尼卫星在飞近这颗卫星时拍的照片，显示了巨大的冰粒子组成的羽状物，在南极区域连续喷射出来。此外，它的表面被碎片覆盖，看起来像是老虎的条纹，且比周围区域温暖，可能与内部区域相连（图 3.13），这意味着这颗卫星内部存在有高压液态水，可能部分与炽热的岩石接触（图 3.14）。卡西尼-INMS 在羽状物中已经检测出双氮、N_2 和（或）CO 及一些碳氢化合物（甲烷、丙烷和乙炔）。N_2 可能由氨的热分解产生，最终由矿物催化，因此，氨及有机物可能

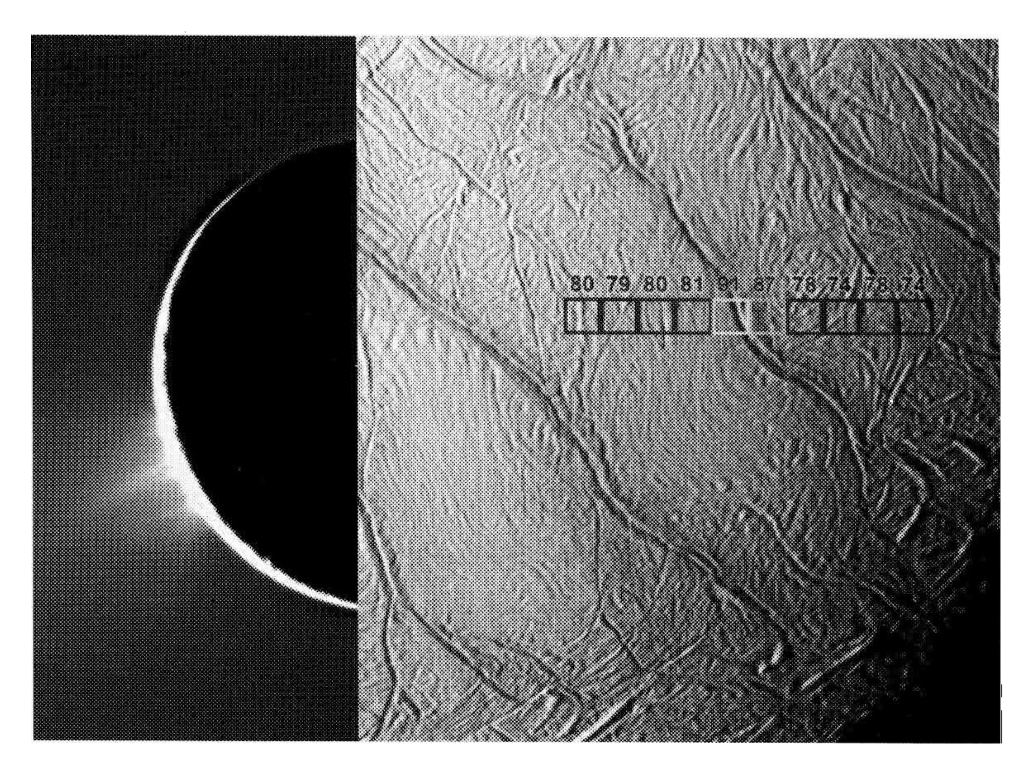


图 3.13　右图：如卡西尼相机所见，在土卫二的南极区域，巨大的羽状物喷射出细小粒子（左图），类似巨大的地球间歇泉。在同一区域，CIRS 仪探测到了更温暖的区域（89～91 K，与周围 74～81 K 的温度相比较），与看起来像“虎纹”（右图）的断裂相关（来源：左图 NASA/JPL/太空科学研究所；右图 NASA/JPL/GSFC/太空科学研究所，见彩插）。

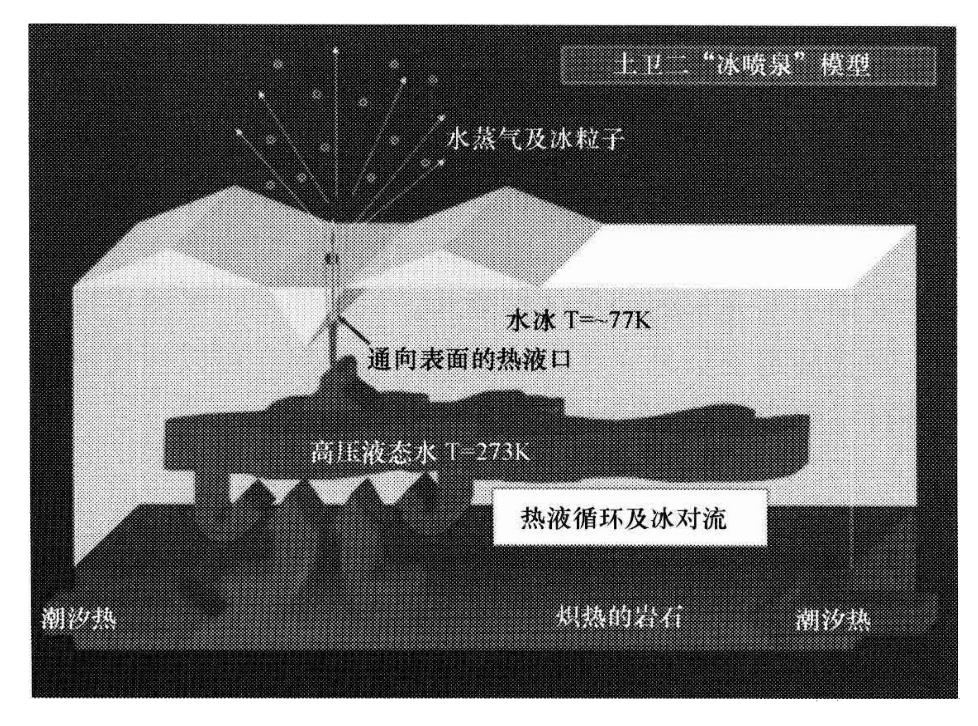


图 3.14　左图。土卫二羽状形成的模型，显示其表面储存有增压的液态水，与高温的岩石物质接触（来源：NASA/JPL/太空科学研究所，见彩插）。

存在于内部的液态水体中，提供了发展前生物化学反应的可能性。实际上，这种环境条件对合成各种有机成分非常有利，包括许多构成生命基础的分子。

3.4 结论：从地球到外太阳系

除地球外，太阳系内仍有一些地方是具有太空生物学意义的。火星仍是探索过去甚至现在生命痕迹的最有吸引力的地点，但外太阳系中还有几个地点，对寻找前生物进程或生命信号有重大意义，木卫二、土卫六，甚至土卫二都是重要的研究对象。然而，就我们现在所知道的可居住星体，很可能地球是太阳系中唯一存在显微可见生命的地点。

现在，对太阳系外行星的探索，使寻找地球外处于更高级进化阶段的生命有了新的可能。10 年内，已经探测了约 300 个系外行星，由于手段非常有限，只能探测非常类似地球的大行星。已经发现了未预料到的大行星，运行轨道与其恒星非常近，因而温度很高。其中一些系外行星可能整体覆盖有热的液态水的海洋，这与已知的水生生命相适应。一些类似地球的系外行星也被探测，但它们比地球大（约两倍），且似乎位于不可居住的区域内。虽然如此，我们应很快就能探测大小和一般条件都更与地球类似的系外行星。

但是，我们怎么能在这些系外行星上找到生命呢？对大气中生命信号的探测是一种可能，但需要光谱学的工具，对雄心勃勃的太空任务（如达尔文任务），这些工具可能在 10 年或几十年内都无法应用。但是，有一个方法已经可以利用，即使是纯理论的，即 SETI（对外星智能的探索）方法。即使生命发展到智能生命，并且达到具有先进科技的文明的可能性非常小（这点我们不知道），但不能排除发生在地球上的这一切也在太阳系外其他可居住的星球上发生，尤其当这样的星球数量非常大的时候（这点我们也不知道）。对这样的文明的年龄和寿命有一个非常强的时间限制。暂时来说，通过辐射信号，科技“先进的”地球文明只能被距我们的太阳 100 光年内的星体探测到。尽管如此，虽然它具有高度的探测性和局限性，寻找假定的信号（在无线电谱域内，但亦可以在其他谱域内）目前仍是探测系外行星生命活动的唯一方法。

参考文献

扩展阅读

1. Gilmour I, Sephton MA (eds). An introduction to astrobiology. Cambridge: USA. Cambridge University Press, 2004:
2. Horneck G, Reetberg P (eds). Complete course in astrobiology. Germany: Wiley-VCH Pub., Weinheim, 2007:
3. Pudritz RE, Higgs P, Stone J (eds). Planetary systems and the origins of life. Cambridge Astrobiology Series 2007; in press.
4. Schulze Makuch D, Irwin LN. Life in the universe. Expectation and constraints. Springer, Berlin 2004.
5. Encrenaz T, Kallenbach R, Owen TC et al. Theouter planets and their moons. Space sciences series of ISSI. Springer, Berlin 2005.
6. Lorenz R, Mitton J. Lifting Titan's veil- exploring the giant moon of Saturn. Cambridge University Press, Cambridge, UK 2002.

专题文献

7. Sotin C, Prieur D. Jupiter's moon Europa: geology and habitability. In ref. 2 2007; 253-271.
8. Chyba CF, Phillips CB. Europa as an abode of life. Origin Life Evol Biosph 2002; 32:47-68.
9. Melosh HJ, Ekholm AG, Showman AP et al. The temperature of Europa's subsurface water ocean. Icarus 2004; 168:498-502.
10. Hand KP, Chyba C. Empirical constraints on the salinity of the Europan ocean and implications for a thin ice shell. Icarus 2007.
11. Oro J, Squyres SW, Reynolds RT et al. Europa: prospects for an ocean and

exobiological implications. In: Exobiology in solar system exploration. G. Carle, D Schwartz and J Huntington eds. NASA SP 1992; 512:102-125.
12. Reynolds R, Squyres S, Colburn D et al. On the habitability of europa. Icarus 1983; 56:246-254.
13. Chela Flores J. Possible degree of evolution of solar-system microorganisms. In: Exobiology: Matter, energy and information in the origin and evolution of life in the universe, J Chela-Flores and F Raulin Eds., Kluwer 1998; 229-234.
14. Souchez R, Jean Baptist P, Petit JR et al. What is the deepest part of the vostok ice core is telling us? Earth-Science Reviews 2002; 60:131-146.
15. Russell CT. Cassini Huygens Mission: Overview, Objectives and Huygens Instrumentarium. Dordrecht, The Netherlands Kluwer Academic Publishers: 2003.
16. Russell CT. The Cassini-Huygens Mission: Orbiter Remote Sensing Investigations. Dordrecht, The Netherlands, Kluwer Academic Publishers: 2005.
17. Sagan C, Khare BN. Tholins. Organic chemistry of interstellar grains and gas. Nature 1979; 277:102-107.
18. Lebreton JP et al. Huygens descent and landing on Titan: Mission overview and science highlights. Nature 2005; 438:758-764.
19. Niemann HB et al. The abundances of constituents of Titan's atmosphere from the GCMS instrument on the huygens probe. Nature 2005; 438:779-784.
20. Israël G et al. Evidence for the presence of complex organic matter in Titan's aerosols by in situ analysis. Nature 2005; 438:796-799. Complex organic matter in Titan's aerosols? (Reply), Nature 2006; 444:E6-E7.
21. Raulin F. Astrobiologiy and habitability of Titan. Space Science Reviews 2007.
22. Waite H et al. The process of tholin formation in Titan's upper atmosphere. Science 2007; 316:870-875.
23. Fortes AD. Exobiological implications of a possible ammonia-water ocean inside Titan, Icarus 2000; 146:444-452.

4 火星，21世纪太空生物学的目标？

Patrice Coll, Fabien Stalport

4.1 引言

火星是太阳系中从太阳开始的第四颗行星，半径约为地球的一半，质量仅为地球的1/10。因为从地球上看外表呈红色，所以它也被称为“红色星球”，这是由于含有铁（Ⅲ）氧化物（通常更多被称为赤铁矿）。

火星是一颗类地行星（图4.1），有一层薄薄的大气，其主要成分为CO_2，表面特征看起来既有类似月亮的陨石坑，也有类似地球的火山、峡谷、荒漠和极地冰盖。除了这些地理特征，火星的转动周期和季节轮回也与地球类似。

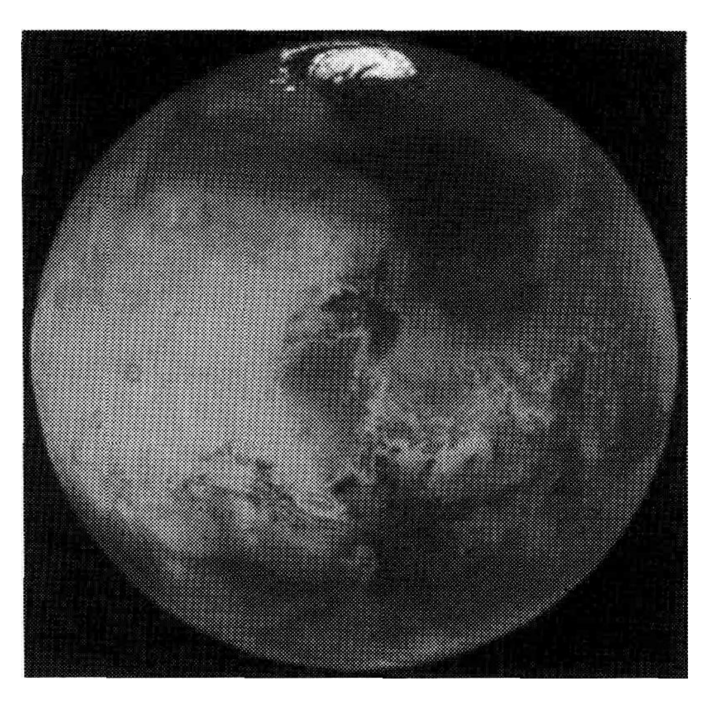


图4.1 火星（来源：NASA，见彩插）。

美国、苏联、欧洲、日本已经发射了许多包括卫星、着陆器和漫游卫星在内的航天器，去研究火星的表面、环境和地质情况。目前，火星是三个功能性轨道航天器的所在地：火星奥德赛号（NASA）、火星特快（ESA）和火星侦查轨道器（NASA），这比除地球外的任何行星都多。它的表面也同时是两个NASA火星探险漫游者（Spirit和Opportunity）的所在地。

所有目前和之前的任务都提供了大量数据，使我们能够更好地理解这颗行星

上发生的理化过程。这里我们将给出一个关于火星研究进展的简短概要，主要来自空间任务，以表明火星是否仍具有太空生物学的吸引力。

4.2 为什么对火星的研究具有太空生物学意义?

火星上过去和（或）现在是否存在生命的问题仍需讨论。在这里提到的第一个重要问题是，在超过 40 年对火星的探测中，只有一个任务具有太空生物学的目的——20 世纪 70 年代中期的 Viking 任务[6,7]。

在两个着陆点，Viking 着陆器收集了生命活动的证据，进行不同测量，给出了正面答案，似乎证明了火星生命的存在。但是，所有这些结果后来都被解释为，由于火星大气/地表界面存在氧化剂而引起的氧化过程[8~15]。更令人惊异的是，GC-MS 实验没有探测到任何有机分子（就我们所知生命必需的一类分子），这似乎与生命的存在相违背[11]。20 世纪 70 年代末，这个任务的结论是没有生命存在或曾经存在于火星上，接着就放弃了对生命活动的寻找。后来在 90 年代，科学家们研究了一颗被认为来源于火星的陨石——ALH84001。与生命过程相关的不同分子被检出，并被认为可能是火星生命[16]，虽然它们的存在也可以被解释为化学过程或地球污染，这些发现突然间又一次激起了人们寻找火星生命的兴趣。从这些研究开始，对火星进行探测的紧张局面到今天仍在继续，揭示出火星在历史上具有产生生命的条件，或至少能够存在前生物化学过程（见 4.3 节）。

那么，回到 Viking 基于太空生物学目的进行的观测，它不能给出任何确凿证据以证明火星上过去或现在具有生命活动，但科学家们认为，也没有明确证据能证明火星上没有生命存在。为什么 Viking 进行的热解-GC-MS 实验没有检测到有机物呢[8,17,18]? 这可能是因为实验条件不适合氨基酸（由于高温而变性）的检测，并且对它们的分解产物（胺类）有一个相对高的检测局限。而且，Viking 着陆器只采集了表面上深度为 1cm 的火星土壤，那里的有机物可能已经被大气中的 H_2O_2[19,20]、原子氧[21]和超氧离子[22]氧化或破坏。Benner[23]表明，它们可能被降解为高度耐热的苯甲酸盐。由于 Viking 的加热器不能够热解这样的分子，不会得到任何预期信号。Viking 着陆器上进行的“生物学”实验，本来是要回答火星上生命存在与否的问题，但其给出的部分结果仍有争议[24~27]。然而，20 世纪 70 年代末，对这些实验结果最悲观的判断是，在火星表面探测过的几平方米内，距离土壤表面 10cm 的深度中，都没有发现类似地球的生命活动存在。

最近，在火星大气中检测到了甲烷[28,29]。火星大气被认为处于热力学平衡中，对这样一个星体，甲烷的存在首先是让人吃惊的[30]。举例来说，地球大气中的甲烷，被认为是不平衡的，直接反映了生命的活动。即使许多无机物可能是火星上甲烷的来源，小生境的存在仍不失为一种合理的可能。不论怎样，最近指

出，甲烷也可能由一个被称为蛇纹石化作用的非生物过程产生，其中有水、二氧化碳和火星上常见的橄榄石的参与[31]。

值得一提的是，根据我们对生命产生过程的了解，火星可能含有这个过程的关键成分。尽管火星和地球的历史进化过程在很久以前就不一样了，但在它们形成后的几亿年内，这两个星球的条件完全有可能相似。来自于35亿年前最古老的沉积物记录告诉我们，这些条件带来了地球最初生命形式的产生[32~35]。所以，如果可再现的理化过程带来的结果是地球上生命的诞生，没有理由认为，同样的过程不能在火星上产生生命。相反，如果能够证明，即使所有条件都合适，火星上也没有出现生命，那么以下问题会很有趣：为什么在地球，而不是火星上产生了生命？

是不是我们对生命必要条件的认识还不完整、不正确？如果这样，缺少的是什么呢？

我们必须时刻记住，火星的地质记录保存得比地球好很多：其表面的一半比38亿年还要久远[36~39]，而地球只有1%的表面是20亿年前的，而板块构造运动使生命起源的前生物印迹都无法挽回地消除了。由于以下两点原因，生命甚至可能首先在火星上出现。

（1）火星的小体积（火星的直径为6800km，约为地球的50%）在最初可能是一个优势，其表面会比地球冷却得更快，会更早产生液态水。

（2）火星没有经历巨大的冲击，不像地球，一次撞击导致了月亮的诞生，并且吹走了地球的原始大气[40,41]。

总之，火星上有无潜在生命活动的问题是基于一定的假设，同时也伴随着理由充分的论点，使我们继续保持兴趣。很明显火星仍是太阳系中太空生物学研究的主要目标：它是除地球外唯一可能发现过去或现在生命痕迹的地点。

4.3 一个过去适宜生命的环境？

现今，没有任何一个已知的理化过程可以明确解释地球上生命的出现。在太阳系中没有任何有意义的比较点，我们不幸遗漏了有关无生命的化学组分向最简单的有机生命体转化过程中所必需的条件。另外，地球的地质学记录没有提供任何有关这种转化的直接线索。地球也没有为我们提供确切的参考证据，例如，若在与原始地球环境条件不同的情况下，生命还能不能出现？但是，我们仍对地球生命出现的必需条件感兴趣：在一定时期内存在液态水，存在有机物质和至少一种可利用的能源。我们将讨论火星表面的地质记录，尤其是它南半球那些最古老的地质表面。

4.3.1 火星上的液态水

4.3.1.1 河流的流动

在1972年，Mariner 9探测器提供的火星表面照片显示了一些地形结构[42,43]，这些结构指向了一个流出峡谷的古老河床，现已干涸。4年后，火星轨道上的Viking探测器拍摄了新的火星表面照片（图4.2），证明了这些地形结构的存在，它们的形成可能意味着存在有强劲且短暂的液态水流[44,45]。

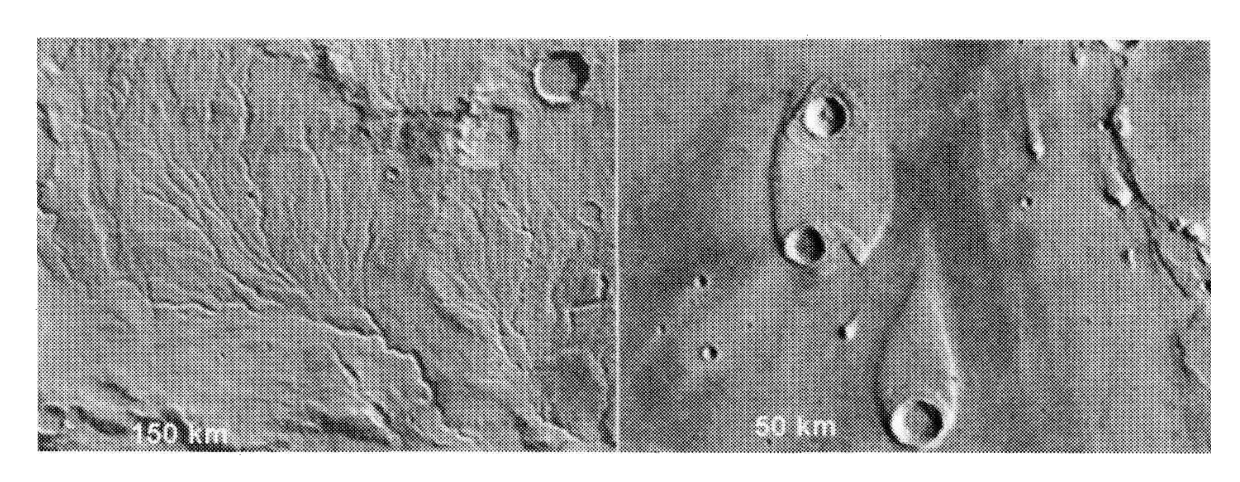


图4.2 火星表面的图像，由Viking探测器在环绕火星的轨道上拍摄。左图，Nirgal区域"树枝状的峡谷"。右图，"外流峡谷"。这些地形结构似乎是在灾难性的外流之后形成（来源：NASA和参考文献45）。

最近，火星奥德赛号上的THEMIS相机（热发射成像系统）使我们能够识别高地上的树状峡谷及水手谷地区的峡谷。这些峡谷的地形特征，尤其是那些与地球上网络结构类似的高密度树状分枝的形成是因为大气降水（雨）带来的水流（图4.3）[46]。通道和网状结构的成熟也许意味着液态水的流动在相当长的地质时期内是持久的，也就是说火星上的水循环已经存在了超过30亿年。

4.3.1.2 沉积物堆积

火星表面，尤其是南半球，布满了陨石撞击的痕迹。这说明它们中的一部分可能暂时被湖占据。盆状的陨石坑可能有利于蓄水池的形成和沉积物的积累。火星表面的照片显示，在一些陨石坑的底部，有明暗地层形成[47~49]。

基于与地球的形态学比较，这种地层的排列常常具有湖中沉积物堆积的特征。因此陨石坑底部的这种层化作用被认为是火星上过去有湖存在的证据（图4.4）。一些沉积物具有"三角洲"的形态，这表示河流对湖进行了液态水的补充。但是，也有其他与液态水无关的假设来解释这些结构的形成：火山熔岩和（或）灰，或者是被风带来的尘埃的堆积[50]。

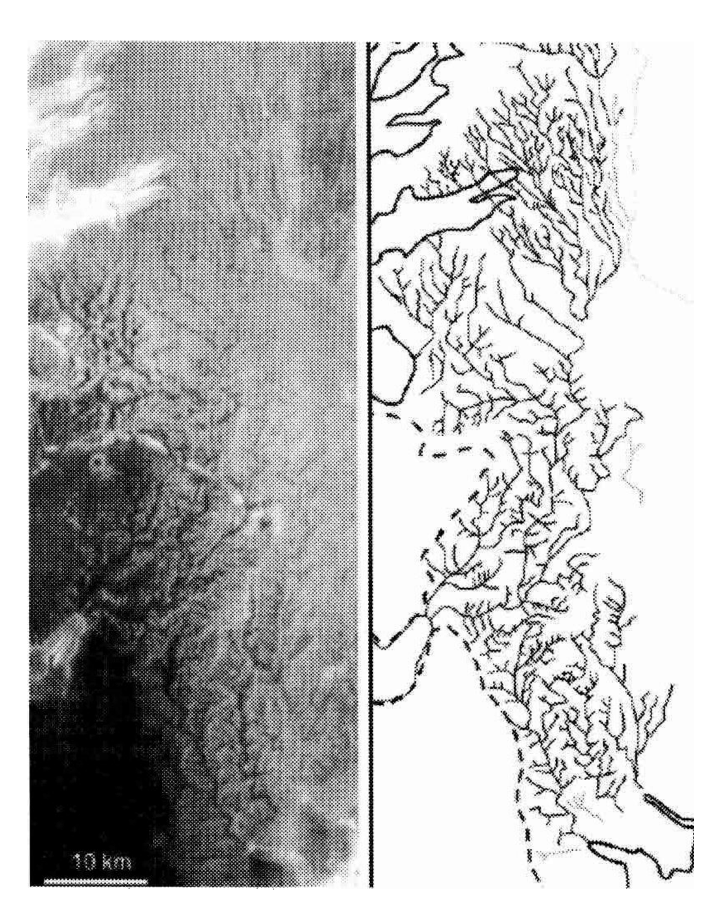


图 4.3　左图为 Valles Marineris 北部的 Fallen Chasma 地区的热红外图像，来自奥德赛号探测器上机载的 THEMIS 仪；右图为河流网络的示意图，来自 Mangold, N.，Quantin, C.，Ansan, V.，Delacourt, C. 和 Allemand, P.，火星上沉淀物的证据来自 Valles Marineris 地区的 Dendritic Valleys。*Science* 2004；305：78-81。经 AAAS 许可再版。

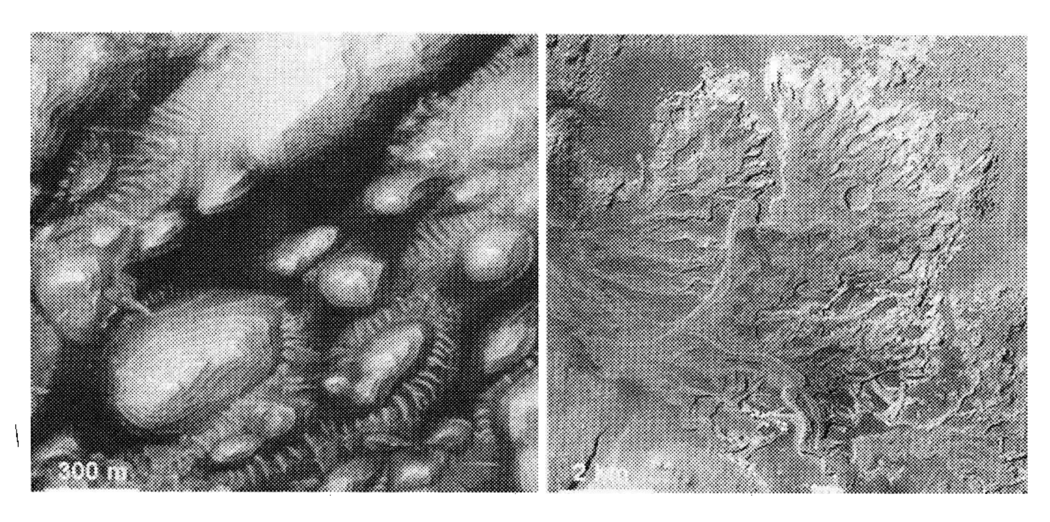


图 4.4　左图，沉积物沉积于 Arabia Terra 地区的 Holden 火山口底部；右图，冲积物形成三角洲。这些结构似乎指示出一个湖畔沉积的环境，可能意味着液态水的存在（来源：参考文献 49）。

2004 年，火星探险漫游者号上的两个漫游者之一——Opportunity 上的全景相机 Pancam，拍下了地球之外的行星上第一张沉积物堆积的照片。在距离这个着陆探测器（图 4.5）几米范围内，沉积物平齐排列，这是在 Terra Maridianni 地区，一个称为 Eagle[51] 的陨石坑内发现的。这个在火星表面第一次对堆积沉积物的发现，是解决液态水存在问题的关键，这些沉积物与那些形态上意味着液态水存在的层化作用对应得非常好。



图 4.5 最早的彩色照片之一，由 Opportunity 漫游者上的全景相机拍摄。这张图显示了沉积物的沉积与 Eagle 火山口的内部坡度齐平，这是 Opportunity 登陆地点的底部。漫游者上的仪器主要探测硫酸盐的存在，它们将是水环境的证据（来源：NASA）。

4.3.1.3 古老的海洋？

一个常被探讨的理论认为，火星表面的低地被大量的液态水占据[52~56]。根据这个理论，在北半球的广大地区，应当暂时形成过一个古老的液态水或泥浆的海洋，称为 Oceanus Borealis，这片海应有数千米深[57]。这个理论还认为，在南半球宽广的撞击盆地 Hellas Planitia 和 Argyre Planitia 中，应当形成了内陆海[58,59]。

火星全局探测者号上的雷达高度计——MOLA，提供了火星表面地形的信息。这台仪器显示，在与海洋可能边界对应的地区（图 4.6），在数百千米的距离内有一个几乎不变的高度，与地球上类似的结构比较，与海岸相对应。但是，这片地区的矿物学组成被确定为火山岩，这一点令人惊异，因为这么大量的水的存在应更有利于形成沉积物，尤其是碳酸盐，而不是火山岩。

这样，许多地形结构似乎可以说明历史早期火星表面液态水的存在，而且它们也可能意味着，水在几亿年的时间范围内始终存在。

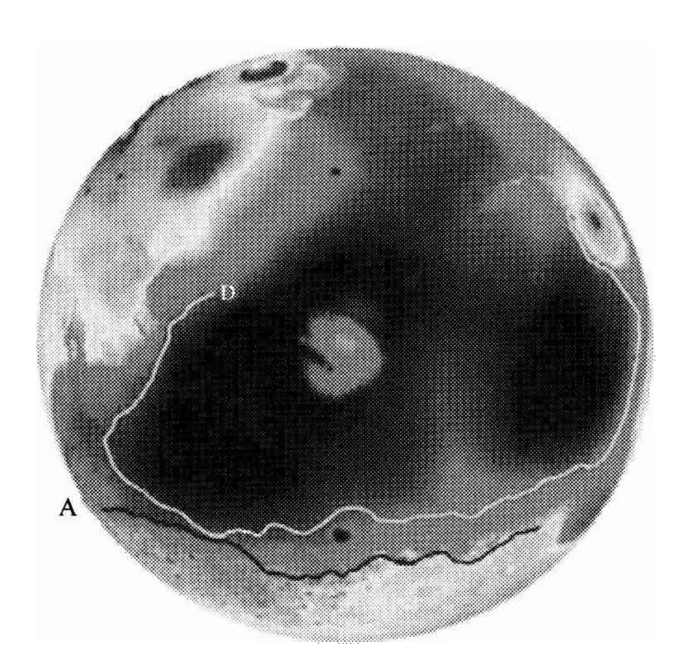


图 4.6　在北半球 Arabia（A）和 Deuteronilus（D）的可能的古代海岸线。经 Macmillan Publishers Ltd 许可再版，*Nature* 447，840～843，版权 2007（见彩插）。

4.3.1.4　矿物学报告

2000 年，火星全局探测者号上的 TES 仪（热发射光谱仪）第一次使我们能够鉴定与液态水存在相一致的矿物痕迹：在如 Sine Meridianni、Valles Marineris 和陨石坑 Aram Chaos 底部（图 4.7）[60,61]等地区发现了灰色或结晶的赤铁矿，在火星尘粒中发现了碳酸盐[62]。但是，这些碳酸盐的来源仍是未知的。

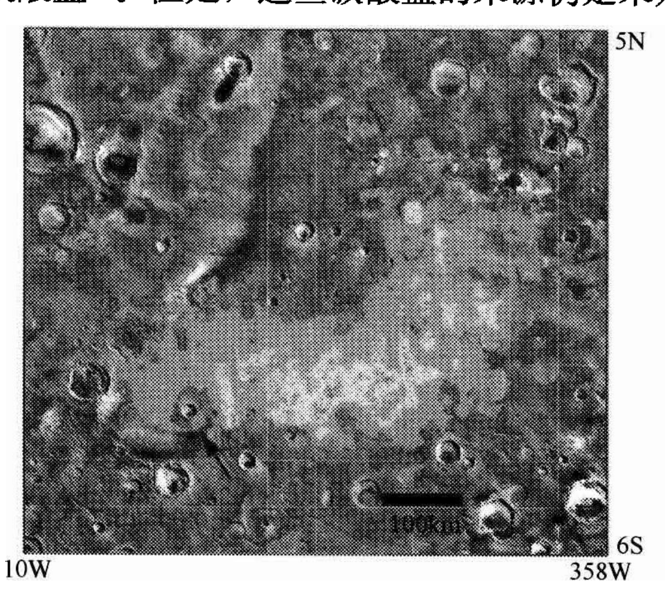


图 4.7　Aram Chaos 火山口的红外图像。在这个火山口的中央，火星全局探测者号上的 TES 仪探测到一个直径 60km 的区域，包含 10%～15%的灰赤铁矿。这个火山口在火星历史的早期可能有一个湖（来源：参考文献 61，见彩插）。

2004 年，NASA 火星探险漫游者（MER）和 ESA 火星特快（MEx），给出了由于液态水存在而形成矿物质的新报告。除了第一次给出沉积物的图像，在两艘 MER 上的仪器还检测出了硫酸盐。在地球上，这些矿物质的特性是只在液态水存在的情况下形成[51,63,64]。

但是，这些硫酸盐的形成可以非常快，这使得水没有必要在较长的地质时期内保持稳定，实际上它们可以在水蒸发的时候沉积下来。火星特快号上的 OMEGA（Observatoire pour la Minéralogie，I' Eau，les Glaces et l' Activité）使我们不仅能检测到硫酸盐[65~67]，而且能检测到黏土的存在[65,68]，这两种物质都具有只在液态水存在时形成的特性。但是，黏土的存在引出了另一个重要的问题：它们的形成可能需要液态水与火星外壳上的硅酸盐岩石进行长地质时期的作用（图 4.8）。最近，火星侦查轨道器上装载的 CRISM（集成型火星侦测成像光谱仪）证实了在 Holden 陨石坑的底部存在有黏土[69]。

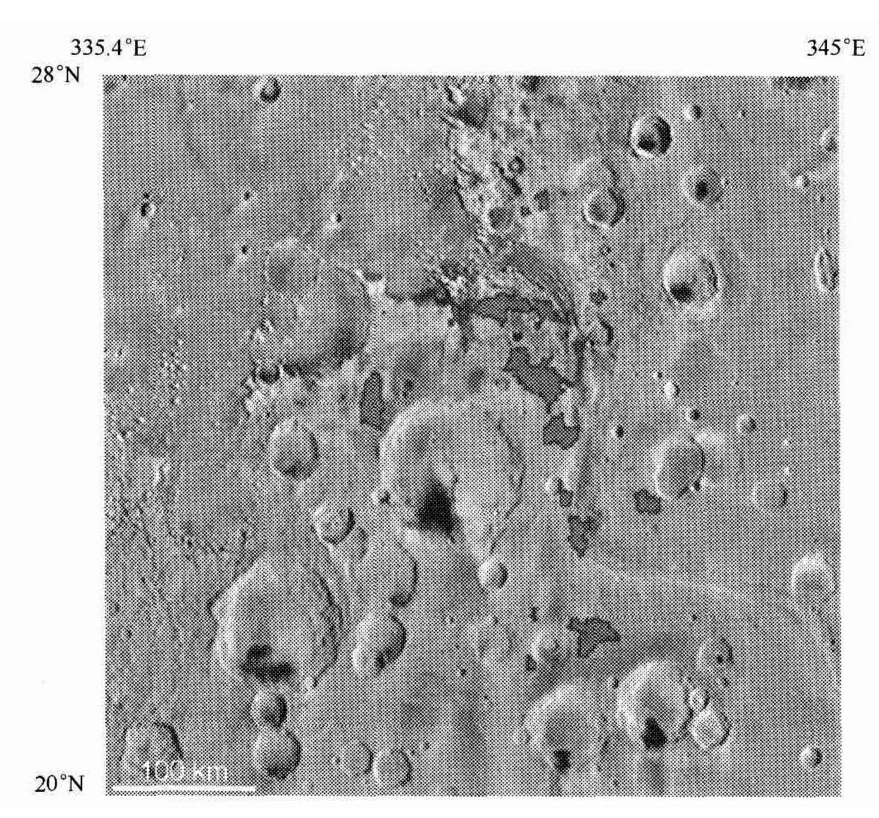


图 4.8 Mawrth Vallis 的红外图像和黏土的鉴定，由火星特快探测器上记载的 OMEGA 仪得到。蓝色为黏土沉积区域。经 Macmillan Publishers Ltd 许可再版：*Nature* 438，623-627，版权 2005（见彩插）。

4.3.1.5 火星上水的历史

考虑到所有这些地形学和矿物学数据，越来越明显的是，在火星已经具备维持水存在的条件后，液态水确实从火星的历史初期就一直存在了。基于火星上水

的历史，几个地质时期应该与两种主要的特定类型的火星岩石——黏土和硫酸盐相对应，因为它们的形成需要液态水存在[70]。在被认为是最古老的表面检测出了黏土，尤其是风力侵蚀或陨星撞击造成的古老沉积面。硫酸盐，即使在火星表面存在了超过 30 亿年，似乎也比黏土更年轻一些。总之，在火星历史上的第一个 7 亿年内，这颗星球可能经历过几次独立的事件，期间液态水量增加或者减少（图 4.9）。

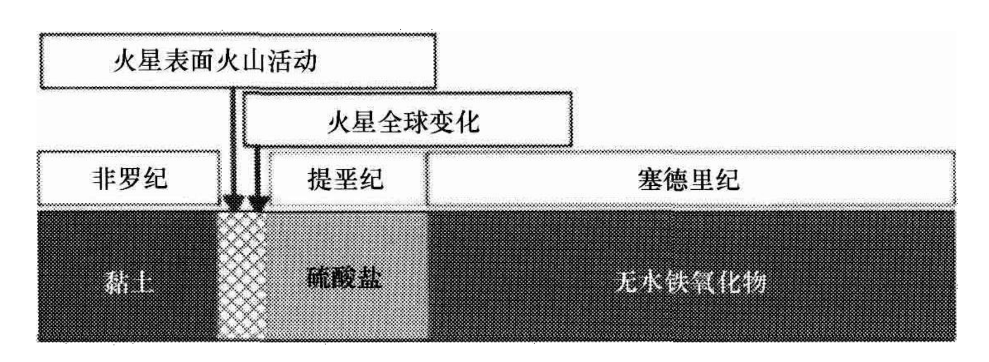


图 4.9　由火星至今的形成历史得到的水的可能历程。最初，水可能含量丰富，并在称为“非罗纪”的时期可能导致了黏土的形成。接着，随着环境改变（即火山活动等等），在“提亚纪”时期，环境变得干燥，酸性更强，有利于硫酸盐沉积。在最近的 38 亿年里，火星环境可能显著恶化，水可能不再以液态形式持久存在，这个时期称为“塞德里纪”。来自 Bibring J-P，Langevin Y，Mustard JF，Poulet FA，Arvidson R，Gendrin A，Gondet B，Mangold N，Pinet P，Forget F 等。火星全球矿物质和水的历程由 OMEGA/火星特快数据推论得到。*Science* 2006；312：100-404。经 AAAS 许可再版。

• 第一个时期，“非罗纪”（phyllosian），期间火星表面液态水含量可能很丰富，可能很深，并且能够在足以形成黏土的时间内改造地表。值得一提的是，这段时期可能相应存在最长为 5 亿年、可使火星大气变薄的磁场[71,72]。

• 第二个时期，“提亚纪”（theiikian），期间环境可能变得干燥，酸性增大，利于硫酸盐的沉积，可能火星表面火山活动活跃[70]。这段时期与磁场消失相对应。

• 第三个时期，“塞德里纪”（siderikian），从 38 亿年前到现在，火星表面的条件衰退得非常严重，液态水消失了（除了一些假设中非常短的灾难性事件[45]或更古老[73]的水流）。接着火星演变到了我们今天所知的环境。这段时期是通过火星表面缓慢的风化作用，导致形成无水的三价铁氧化物。

4.3.2　有机分子的来源

火星表面布满了陨石坑。此外，其表面越古老，陨石坑越多、越大。许多来

自行星际介质的星体撞击了火星表面，尤其是在其早期历史中[36,37,34]。目前，几百吨来自行星际介质的物体落到地球表面[75~78]，也落到火星表面。总的来说，自形成以来，火星表面极有可能已经接受了来自行星际介质的有机物质（图4.10）。内源的有机合成也有可能已经发生，尤其是在大气中[87~90]。火星表面可能已经有许多有机物来源，尤其在它最初的几亿年里。

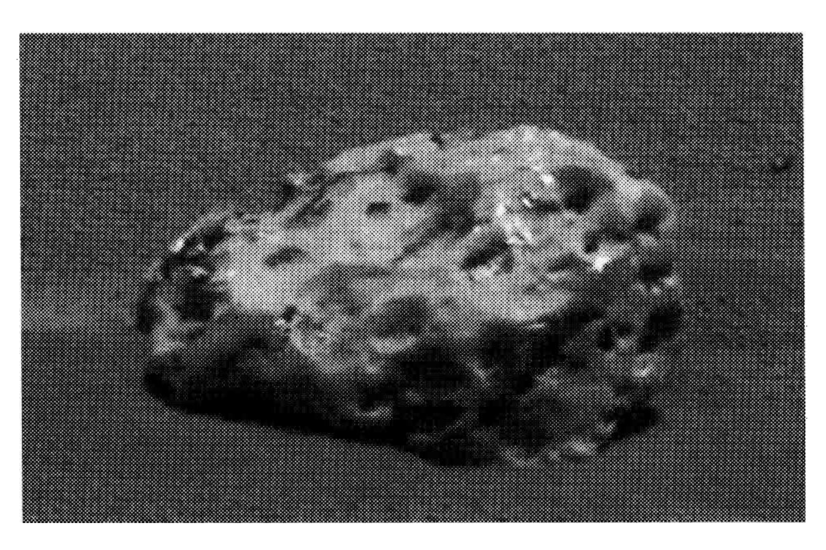


图4.10　在火星表面鉴定的第一块陨石的照片，即在除地球之外的行星表面。这块陨石由Opportunity漫游者上的全景相机在接近其挡热板时定位，因此被命名为“挡热板岩石”，其铁和镍的化学组成与地球上发现的铁陨石非常类似（来源：NASA/JPL/Cornell）。

4.3.3　能量来源

围绕火星上水的历史，可能仍有许多不确定性，但液态水的存在，尤其在黏土形成期间，极有可能需要足够温和的环境条件。这种环境的存在，只有当火星早期曾处于可居住区时才有可能[91~93]。它到太阳的距离，以及45亿年前年轻的太阳较弱的发光能力（比今天少约25%～30%的能量），可能共同降低了液态水存在的可能性。实际上，在没有大气的情况下，火星的平均恒定温度应约为－75℃，因此，火星可能从许多其他能源获利，来达到有利的环境条件：

（1）有一个高密度的大气，主要由温室气体[97,98]和（或）碳冰形成的云组成[99]；

（2）由大量能量产生的地热流，其能量来源于：①冲击层储存；②放射性元素衰变；③ 液态核不断结晶化释放的潜在热量；

（3）许多撞击火星表面的太阳系星体可能暂时或局部加热了火星表面。这些星体的撞击可能为岩石熔化等提供了能量。

4.3.4 关于潜在生物活动的结论

在火星历史上第一个7亿年间，其表面就具备了生命活动产生和维持的所有必要条件。尽管导致地球上第一个活生物诞生的自发化学反应仍然未知，而且与地球类似的环境条件未必就能产生生命，以下问题仍很合理：既然在相同时期火星存在着与地球相似的环境条件，生命是否也能够在火星上产生?

4.4 去哪儿寻找及怎样寻找生命?

火星遥远的过去决定了在不久的将来人们如何对这颗红色星球进行探索。欧洲的火星特快和美国的Spirit和Opportunity漫游者最近对火星的探索，将使研究重新定向于火星最古老表面逝去的生命。陨星的撞击和风类似于水流的侵蚀，可能猛烈而短暂，使得埋藏的地层得以暴露，古代发生的事件得以被探知，那时火星上可能还存在有液态水。目前，考虑到这个任务的太空生物学目的，科学界正在确定MSL漫游者的着陆地点。在候选位置中，由摄谱成像仪OMEGA探测到的黏土可能性最大。在火星环境变得不良的数十亿年里，这类地表可能将早期生命痕迹以化石的形式保护起来。

2009年，NASA发射MSL（火星科学实验室）漫游者到火星，这次任务的科学目的主要是寻找已灭绝或现存的生命痕迹，或可能存在的前生物化学反应。它将使用钻探系统，来达到几分米的深度，最终达到采样区，在那里无机物可能经历了火星大气还能够养育生命的时代，有机残留物可能从毁灭中保留了下来[100,101]。

MSL漫游者机载的SAM实验是为了对火星大气、地面及地下进行分析。

（1）对重新取得的地面或地下样本的元素组成进行确认，以与其他着陆器上的实验结果对比，并且为以下（2）、（6）、（7）条提供比较点。

（2）在地面或地下，寻找可能经历过过去前生物化学反应的有机物痕迹，在最好的情况下，寻找可能在火星表面或地下产生过的有机生命体。有机物分子的手性可以提供有关它们来源的线索。与对有机分子的寻找相平行的是对同位素比例的研究，它可以提供更多信息。

（3）获取大气中稀有气体（Ar、Xe、Kr）同位素含量和比例的信息，CO_2中的C比（13/12）、N_2中的N比（15/14）和H_2O中的D/H比。这将能够与其他观察（来自着陆器及地球上）相对应，使我们更好地理解火星历史上地壳的排气和大气的逃逸。

（4）在对地表（或地下）样本进行激光消融或加热后，测定其碳、氮、硫（34/32）、氧（16/17/18）和D/H的同位素比例，与来自地球的测量结果相关联，并与对其大气的测量结果相比较。

（5）寻找大气中与（2）条相关，来源可能为有机物的气体（CH_4、H_2S等）

（6）获取以下项目存在的信息：地表（或地下）样本中氢氧化物、碳酸盐、硫酸盐及黏土；在浓密及湿润的大气中的矿物指标，它将对在地面寻找有机物痕迹有所帮助［（2）条］。并且可以通过简化的特定热分析将生物矿物中的矿物质区分开来。

（7）若钻探系统能够获得不同深度的样本，不仅对上述的（1）、（2）、（4）、（6）条，也可由吸收的固相气体（H_2O、CO_2、H_2O_2等）获取来自不同地层的信息。

4.5　正在进行的探索

目前对行星可居住性（产生及维持生命的能力）的理解，主要集中在那些表面具有液态水的行星。不同方面的证据表明，火星过去的可居住性要明显高于现在，但仍不知道那里是否曾经存在过活的生命体。20世纪70年代中期，Viking探测器在各自的着陆点进行了实验，探测火星土壤中的微生物，发现暴露于水和营养物之后，它的CO_2产量有暂时的增加。但是，许多科学家随后对“生命”的迹象表示出质疑。以现代关于嗜极生命的知识，对距现在已有30年的Viking数据进行再分析，发现Viking实验对探测这样的生命形式不够精细。

NASA凤凰号火星登录器，发射于2007年8月4号，在2008年5月25号成功到达火星北极地区的“绿谷”，在几周之后的7月31号，用其热及演化气体分析仪或TEGA，成功对火星土壤的第一份样本进行了分析。通过对样本进行加热，在样本温度达到0℃时，质谱仪检测到了水蒸气。“这里有水”，TEGA的首席科学家William Boynton总结说，“以前，来自火星奥德赛号观测到的水冰和上个月凤凰号观测到的消失的冰块，我们已经见过一些证据，但这次是我们第一次接触并体验到了火星上的水。”

继这次来自凤凰号的历史性的发现，对火星的探测将会继续，2009年，通过火星科学实验室（图4.11），火星探险漫游者更大、更快、更强大的版本，以

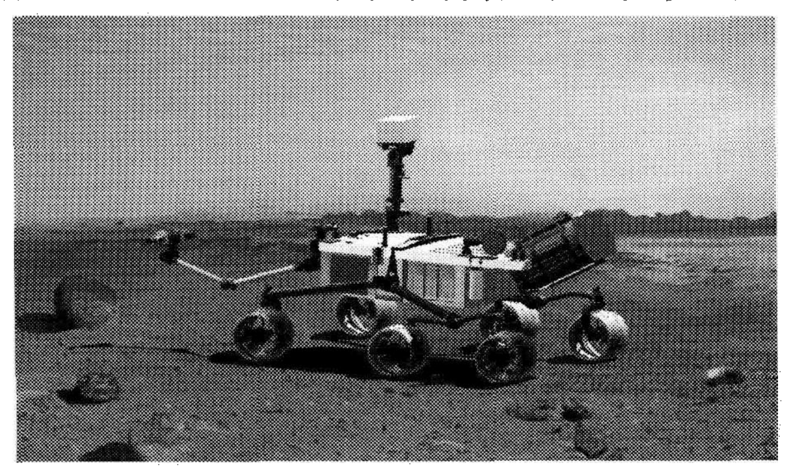


图4.11　MSL 2009的景观（来源：NASA MSL网站）。

及2012年ESA的第一个火星漫游者，将能够钻探到土壤2 m深的地方，以寻找有机分子。这些及以后的任务将使我们离这个关键问题的答案更近："火星上曾经存在过生命吗?"

参考文献

扩展阅读

1. Kieffer HH, Snyder C, Jakosky BM et al. Mars. University of Arizona Press, 1977.
   - 仍然是最好的参考书，有1500页及114个作者。
2. Forget F, Costard F, Lognonné P. La planète Mars—Histoire d'un autre monde, Belin, 2006.
   - 若您能阅读法文的话。
3. Rocard F. Dernière nouvelles de Mars. Planète Rouge. Dunod, 2006.
   - 若您能阅读法文的话。
4. Zubrin R, Wagner R. The case for Mars. Simon and Schuster Trade, 1996.
   - 若您有兴趣看到人类到访火星的话。
5. Jakosky B. The search for life on other planets. Cambridge University Press, 1998.
   - 若您对前往其他行星寻找生命有兴趣，亦包括火星，是最好的一本书！

专题文献

6. Soffen GA, Snyder CW. The first Viking mission to Mars. Science 1976; 193:759-766.
7. Soffen GA. The Viking project. J Geophys Res 1977; 82:3959-3970.
8. Biemann K, Oro J, Toulmin P III et al. Search for organic and volatile inorganic compounds in two surface samples from the Chryse Planitia region of Mars. Science 1976; 194:72-76.
9. Horowitz NH, Hobby GL, Hubbard JS. The Viking carbon assimilation experiments—Interim report. Science 1976; 194:1321.
10. Klein HP, Horowitz NH, Levin GV et al. The Viking biological investigation: preliminary results. Science 1976; 194:99-105.
11. Biemann K, Oro J, Toulmin P et al. The search for organic substances and inorganic volatile compounds in the surface of Mars. J Geophys Res 1977; 82:4641-4658.
12. Klein HP. The Viking biological investigation: general aspects. J Geophys Res 1977; 82:4677-4680.
13. Levin GV, Straat PA. Recent results from the Viking Labeled Release experiment on Mars. J Geophys Res 1977; 82:4663-4667.
14. Oyama VI, Berdahl BJ. The Viking gas exchange experiment results from Chryse and Utopia surface samples. J Geophys Res 1977; 82:4669-4676.
15. Oyama VI, Berdahl BJ, Carle GC. Preliminary findings of the Viking gas exchange experiment and a model for martian surface chemistry. Nature 1977; 265:110-114.
16. McKay DS, Gibson EK, Thomas-Keprta KL et al. Search for past life on Mars: Possible relic biogenic activity in martian meteorite ALH84001. Science 1996; 273:924-930.
17. Klein HP, Horowitz NH, Biemann K. The search for extant life on Mars. In: Kieffer HH, Jakosky BM, Snyder C, Matthews MS, eds Mars. University of Arizona Press, Tucson 1992; 1221-1234.
18. Glavin DP, Schubert M, Botta O et al. Detecting pyrolysis products from bacteria on Mars. Earth Planet Sci Lett 2001; 185:1-5.
19. Bullock MA, Stoker CR, McKay CP et al. A coupled soil-atmosphere model of H_2O_2 on Mars. Icarus 1994; 107:142-154.
20. Zent AP. On the thickness of the oxidized layer of the martian regolith. J Geophys Res 1998; 103:31491-31498.
21. Kolb C, Lammer H, Abart R et al. The martian oxygen surface sink and its implications for the oxidant extinction depth. In: Proceedings of the Second European Workshop on Exo/Astrobiology 16-19 September 2002, Graz, Austria. Ed.: Huguette Lacoste. ESA SP-518, Noordwijk, Netherlands: ESA Publications Division 181-186.
22. Yen AS, Kim SS, Hecht MH et al. Evidence that the reactivity of the martian soil is due to superoxide ions. Science 2002; 289:1909-1912.
23. Benner SA, Devine KG, Matveeva LN et al. The missing organic molecules on Mars. PNAS 2000; 97:2425-2430.
24. Navarro-Gonzalez R, Navarro KF, Rosa Jdl et al. The limitations on organic detection in Mars-like soils by thermal volatilization-gas chromatography-MS and their implications for the Viking results. PNAS 2006; 103:16089-16094.
25. Biemann K. On the ability of the Viking gas chromatograph-mass spectrometer to detect organic matter. PNAS 2007; 104:10310-10313.
26. Mukhopadhyay R. The Viking GC/MS and the search for organics on Mars. Anal Chem 2007; 79:7249-7256.
27. Wu C. Secrets of the martian soil. Nature 2007; 448:742.
28. Formisano V, Atreya S, Encrenaz T et al. Detection of methane in the atmosphere of Mars. Science 2004; 306:1758-1761.
29. Krasnopolsky VA, Maillard JP, Owen TC. Detection of methane in the martian atmosphere: evidence for life? Icarus 2004; 172:537-547.
30. Lewis JS, Prinn RG. Planets and their atmospheres—origin and evolution. Publication by Orlando FL Academic Press Inc International Geophysics Series 1984; 33:480.
31. Oze C, Sharma M. Have olivine, will gas: Serpentinization and the abiogenic production of methane on Mars. Geophys Res Lett 2005; 32:10203.
32. Lowe DR. Stromatolites 3,400-Myr old from the Archean of Western Australia. Nature 1980; 284:441-443.
33. Walter MR, Buick R, Dunlop JSR. Stromatolites 3,400-3,500 Myr old from the North Pole area, Western Australia. Nature 1980; 284:443-445.
34. Byerly GR, Lower DR, Walsh MM. Stromatolites from the 3,300-3,500-Myr Swaziland Supergroup, Barberton Mountain Land, South Africa. Nature 1986; 319:489-491.
35. Allwood AC, Walter MR, Kamber BS et al. Stromatolite reef from the early archaean era of Australia. Nature 2006; 441:714-718.
36. Hartmann WK, Neukum G. Cratering chronology and the evolution of Mars. Space Sci Rev 2001; 96:165-194.
37. Neukum G, Ivanov BA, Hartmann WK. Cratering records in the inner solar system in relation to the lunar reference system. Space Sci Rev 2001; 96:55-86.
38. Frey HV, Roark JH, Shockey KM et al. Ancient lowlands on Mars. Geophys Res Lett 2002; 29:1384.
39. Frey HV. Impact constraints on the age and origin of the lowlands of Mars. Geophys Res Lett 2006; 33, doi:10.1029/2005GL024484.
40. Hartmann WK. Lunar origin: The role of giant impacts. Abstracts and program for the conference on the origin of the Moon Kona Hawaii. LPI Contribution 540, published by the Lunar and Planetary Institute 1984; 540:52.
41. Dickinson T, Newsom HE. A possible test of the impact theory for the origin of the moon. Lunar and planetary science XVI 1985; 183-184.
42. Masursky H. An overview of geological results from Mariner 9. J Geophys Res 1973; 78:4009-4030.
43 Baker VR, Milton DJ. Erosion by catastrophic floods on Mars and Earth. Icarus 1974; 23:27-41.
44. Carr MH. Formation of martian flood features by release of water from confined aquifers. J Geophys Res 1979; 84:2995-3007.
45. Baker VR, Carr MH, Gulick VC et al. In: Kieffer HH, Jakosky BM, Snyder C, Matthews MS Eds., Mars, University of Arizona Press, Tucson 1992:493-522.
46. Mangold N, Quantin C, Ansan V et al. Evidence for precipitation on Mars from dendritic valleys in the valles marineris area. Science 2004; 305:78-81.
47. Edgett KS, Parker TJ. Water on early Mars: Possible subaqueous sedimentary deposits covering ancient cratered terrain in western Arabia and Sinus Meridiani. Geophys Res Lett 1997; 24:2897.
48. Moore JM, Howard AD, Dietrich WE et al. Martian layered fluvial deposits: Implications for Noachian Climate Scenarios. Geophys Res Lett 2003; doi:10.1029/2003GL019002.
49. Moore JM, Howard AD. Large alluvial fans on Mars. J Geophys Res (Planets) 2005; 110:E04005 doi:10.1029/2004JE002352.
50. Leverington DW, Maxwell TA. An igneous origin for features of a candidate crater-lake system in western Memnonia, Mars. J Geophys Res (Planets) 2004; 109:E06006 doi:10.1029/2004JE002237.
51. Squyres SW, Arvidson RE, Bell JF et al. The Opportunity Rover's Athena science investigation at Meridiani Planum, Mars. Science 2004; 306:1698-1703.
52. Head JW, Kreslavsky M, Hiesinger H et al. Oceans in the past history of Mars: Tests for their presence using Mars Orbiter Laser Altimeter (MOLA) data. Geophys Res Lett 1998; 25: 4401-4404.
53. Parker TJ. Mapping of possible "Oceanus Borealis" shorelines on Mars:

A status report. 29th annual lunar and planetary science conference, Houston, TX 1998, abstract no. 29; 1965.
54. Clifford SM, Parker TJ. The evolution of the martian hydrosphere: Implications for the fate of a primordial ocean and the current state of the northern plains. Icarus 2001; 154:40-79.
55. Sotin C, Couturier F, Bibring J. Analysis of gravity potential along paleo-shorelines on Mars: Implications for ocean on very early Mars. Eos Trans. AGU, 87(52) Fall Meeting 2006 Suppl, Abstract P31B-0139.
56. Perron JT, Mitrovica JX, Manga M et al. Evidence for an ancient martian ocean in the topography of deformed shorelines. Nature 2007; 447:840-843.
57. Öner AT, Ruiz J, Fairén AG et al. The volume of possible ancient oceanic basins in the northern plains of Mars. Lunar and Planetary Science XXXV 2004; Abstract no. 35: 1319.
58. Head JW. Exploration for standing bodies of water on Mars: When were they there, where did they go and what are the implications for astrobiology? Eos Trans. AGU, 82(47) Fall Meeting 2001 Suppl, Abstract P21C-03.
59. Hiesinger H, Head JW III. Topography and morphology of the Argyre Basin, Mars: Implications for its geologic and hydrologic history. Planet Space Sci 2002; 50:939-981.
60. Christensen PR, Bandfield JL, Clark RN et al. Detection of crystalline hematite mineralization on Mars by the Thermal Emission Spectrometer: Evidence for near-surface water. J Geophys Res 2000; 105:9623-9642.
61. Christensen PR, Morris RV, Lane MD et al. Global mapping of martian hematite mineral deposits: Remnants of water-driven processes on early Mars. J Geophys Res 2001; 106:23873-23886.
62. Bandfield JL, Glotch TD, Christensen PR. Spectroscopic identification of carbonate minerals in the martian dust. Science 2003; 301:1084-1087.
63. Christensen PR, Wyatt MB, Glotch TD et al. Mineralogy at Meridiani Planum from the Mini-TES Experiment on the opportunity rover. Science 2004; 306:1733-1739.
64. Klingelhöfer G, Morris RV, Bernhardt B et al. Jarosite and hematite at Meridiani Planum from Opportunity's Mössbauer Spectrometer. Science 2004; 306:1740-1745.
65. Bibring J-P, Langevin Y, Gendrin A et al. Mars surface diversity as revealed by the OMEGA/Mars Express observations. Science 2005; 307:1576-1581.
66. Gendrin A, Mangold N, Bibring J-P et al. Sulfates in martian layered terrains: The OMEGA/Mars Express view. Science 2005; 307:1587-1591.
67. Langevin Y, Poulet F, Bibring J-P et al. Sulfates in the north polar region of Mars detected by OMEGA/Mars Express. Science 2005; 307:1584-1586.
68. Poulet F, Bibring J-P, Mustard JF et al. Phyllosilicates on Mars and implications for early martian climate. Nature 2005; 438:623-627.
69. Milliken RE, Grotzinger J, Grant J et al. Clay minerals in Holden crater as observed by MRO CRISM. Seventh International Conference on Mars, LPI Contributions 2007, Abstract no 1353:3282.
70. Bibring J-P, Langevin Y, Mustard JF et al. Global mineralogical and aqueous Mars history derived from OMEGA/Mars Express data. Science 2006; 312:100-404.
71. Acuna MH, Connerney JE, Lin RP et al. The magnetic field of Mars—A window into Mars' past. AGU Fall Meeting 2001; Abstracts no. P41A-06.
72. Acuña MH, Connerney JEP, Wasilewski P et al. Magnetic field of Mars: Summary of results from the aerobraking and mapping orbits. J Geophys Res 2001; 106:23403-23418.
73. Dickson JL, Head JW, Marchant DR et al. Recent gully activity on Mars: Clues from late-stage water flow in gully systems and channels in the antarctic dry valleys. Seventh International Conference on Mars 2007; 38:1678.
74. Chapman CR, Cohen BA, Grinspoon DH. What are the real constraints on the existence and magnitude of the late heavy bombardment? Icarus 2007; 189:233-245.
75. Halliday I, Blackwell AT, Griffin AA. The flux of meteorites on the Earth's surface. Meteoritics 1989; 24:173-178.
76. Love SG, Brownlee DE. A Direct measurement of the terrestrial mass accretion rate of cosmic dust. Science 1993; 262:550.
77. Bland PA, Smith TB, Jull AJT et al. The flux of meteorites to the Earth over the last 50,000 years. Mon Not R Astron Soc 1996; 283:551.
78. Taylor S, Lever JH, Harvey RP. Accretion rate of cosmic spherules measured at the South Pole. Nature 1998; 392:899.
79. Clemett SJ, Maechling CR, Zare RN et al. Identification of complex aromatic molecules in individual interplanetary dust particles. Science 1993; 262:721-725.
80. Maurette M, Brack A, Kurat G et al. Were micrometeorites a source of prebiotic molecules on the early Earth? Adv Space Res 1995; 15:113-126.
81. Brinton KLF, Engrand C, Glavin DP et al. A search for extraterrestrial amino acids in carbonaceous antarctic micrometeorites. Orig Life Evol Biosph 1998; 28:413-424.
82. Clemett SJ, Chillier XDF, Gillette S et al. Observation of indigenous polycyclic aromatic hydrocarbons in 'giant' carbonaceous antarctic micrometeorites. Orig Life Evol Biosph 1998; 28:425-448.
83. Maurette M. Carbonaceous micrometeorites and the origin of life. Orig Life Evol Biosph 1998; 28:385-412.
84. Botta O, Bada JL. Extraterrestrial organic compounds in meteorites. Surv Geophys 2002; 23:411-467.
85. Glavin DP, Matrajt G, Bada JL. Re-examination of amino acids in Antarctic micrometeorites. Adv Space Res 2004; 33:106-113.
86. Pizzarello S, Cooper GW, Flynn GJ. In meteorites and the early solar system II. Lauretta DS and McSween HY Eds. Published by The University of Arizona Press. Tucson 2006; 625-651.
87. Chyba C, Sagan C. Cometary and asteroidal delivery of prebiotic organics vs in situ production on the early Earth. Bulletin of the American Astronomical Society 1990; 22:1097.
88. Chyba CF, Thomas PJ, Brookshaw L et al. Cometary delivery of organic molecules to the early Earth. Science 1990; 249:366.
89. Chyba CF, Sagan C. Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: an inventory for the origins of life. Nature 1992; 355:125-132.
90. Bernstein M. Prebiotic materials from on and off the early Earth. Philos Trans R Soc Lond B Biol Sci 2006; 361:1689-1702.
91. Kasting JF, Whitmire DP, Reynolds RT. Habitable zones around main sequence stars. Icarus 1993; 101:108-128.
92. Kasting JF. Habitability of planets. Astrobiology workshop 1996 Abstract.
93. Kasting JF. Habitable zones around stars and the search for extraterrestrial life. Eos Trans. AGU, 82(47), Fall Meet. Suppl, Abstract P21C-01, 2001.
94. Newman MJ, Rood RT. Implications of solar evolution for the Earth's early atmosphere. Science 1977; 198:1035-1037.
95. Gough DO. Solar interior structure and luminosity variations. Solar Phys 1981; 74:21-34.
96. Gilliland RL. Solar evolution. Glob Planet Change 1989; 1:35-55.
97. Kasting JF, Brown LL, Acord JM et al. Was early Mars warmed by ammonia? In: Lunar and Planetary Inst., Workshop on the Martian Surface and Atmosphere Through Time 1992:84-85.
98. Kasting JF. Greenhouse models of early Mars climate. Eos Trans. AGU Fall Meeting 2002 Suppl, Abstracts no. P51B-0345.
99. Forget F, Pierrehumbert RT. Warming early Mars with carbon dioxide clouds that scatter infrared radiation. Science 1997; 278:1273.
100. McKay CP, Friedman EI, Wharton RA et al. History of water on Mars: a biological perspective. Adv Space Res 1992; 12:231-238.
101. Brack A, Clancy P, Fitton B et al. An integrated exobiology package for the search for life on Mars. Adv Space Res 1999; 23:301-308.

5 彗星与太空生物学

Hervé Cottin, Didier Despois

5.1 引言

相当长的时间内，彗星被认为是凶兆，与迷信、战争、死亡、疾病等联系在一起。这种神奇的物体没有预兆地在天空中突然出现又突然消失，关于它并没有真正的了解。亚里士多德认为这是一种大气现象，直到 17 世纪西方科学家才把它们确实看做天体。

1950 年，在一篇里程碑式的论文中，F. Whipple 把彗星描述成“脏雪球”，是冰［主要是水冰（water ice)］和矿物质的混合物[7]。这个模型早已演变，由基于地球的观测、就地观测、样本带回和实验室研究得到的数据，关于其核的性质已有更多了解。到现在，我们已有的知识赋予了彗星在太空生物学中特殊地位：它们是起源的转折点，包括太阳系的起源，因为它们在太阳系增长的地点获得了太阳星云的物质；并且，也有可能包括地球生命的起源。彗星含有丰富的水和碳这两种地球生命所必需的组件，因此地球上部分的水和碳可能来自彗星。

20 世纪早期，Chamberlin 和 Chamberlin 提出，落下的球粒状陨石可能是地球上有机物的重要来源[8]。在 1961 年，根据对含碳和氮的彗星头的观测，J. Oró 第一次提出彗星可能扮演了类似角色：

“我认为彗星与地球相互作用的一个重要结果，是我们星球上相对大量的碳组分的积累，已知它们会自发转变成氨基酸、嘌呤和其他生化组分”。

最近一本书《彗星Ⅱ》给出了一些有关彗星的知识[10]，强烈建议希望更多了解彗星的读者参阅。彗星主要被认为是地球储存有机分子的潜在地点，在这一章，我们首先介绍它们的一般特性，然后介绍彗星物质的化学组成，它们是由观测、就地观测、样本带回和实验室研究推测出来的。接着给出来自 Rosetta 号的预测结果，尤其是来自那些探测彗星环境分子组成的仪器：COSAC，COSIMA，MIRO 和 VIRTIS。最后，介绍彗星对地球可能做出的不同贡献。

5.2 太阳系中的彗星

彗星是太阳系中行星形成后的残余物。它们在第一个 100 万年，前太阳系星

云最冷的部分形成，那的温度低得足以使水冰凝缩，嵌入由有机物和（或）矿物质组成的“尘埃”粒子，形成的“脏雪球”是彗星的核，一般大小为 1～100km：著名的 1P/Halley 彗星的核大小不同寻常，约为 8km×8km×16 km，这是根据 ESA 探测器 Giotto 于 1986 年飞过它时拍摄的照片测量的。

多数彗星都位于两个富集地：① 奥尔特云（Oort cloud），一个巨大的（10^5 AU[a]）球状富集地，彗星数量估计约为 10^{11}～10^{12}，总质量为地球质量的几十分之一；② Kuiper 带，是比海王星轨道更远的一个相当平坦的平面，比奥尔特云小（最大为 100～1000 AU)，含许多 1000 km 级的星体（矮行星），如冥王星和最近发现的 Eris、Sedna 和 Quaoar 跨海王星物体（TNO）。Kuiper 带中的部分天体形成的位置和现在它们到太阳的距离相同，但 Kuiper 带中其他的天体和整个奥尔特云都是在离太阳更近的地方形成，后来由于和大行星（如木星、土星、天王星和海王星）的重力作用，移动到了现在的位置。

我们关于彗核组分的多数资料都是间接的，主要来自地面或来自对它接近太阳时释放的粒子和气体的就地观察。如果平均到整个表面，彗星一个来回会损失物质，将体现为半径减小约 1m。

气体和尘埃粒子以类似球壳的形状扩张开来，某些可抵抗太阳辐射引起的光解和光电离的彗星的头可能会达到 100 万 km。彗尾形成于彗星头部离子与太阳风，以及彗星头部尘埃粒子与太阳辐射压的相互作用（也能观察到一个中性的钠尾）。彗尾可以延长到 1 亿 km，几乎是太阳到地球的距离。虽然比核的可见度高，但彗星头部和彗尾的密度极低，多数地方低于每立方厘米 10^4 原子（图 5.1和图 5.2）。

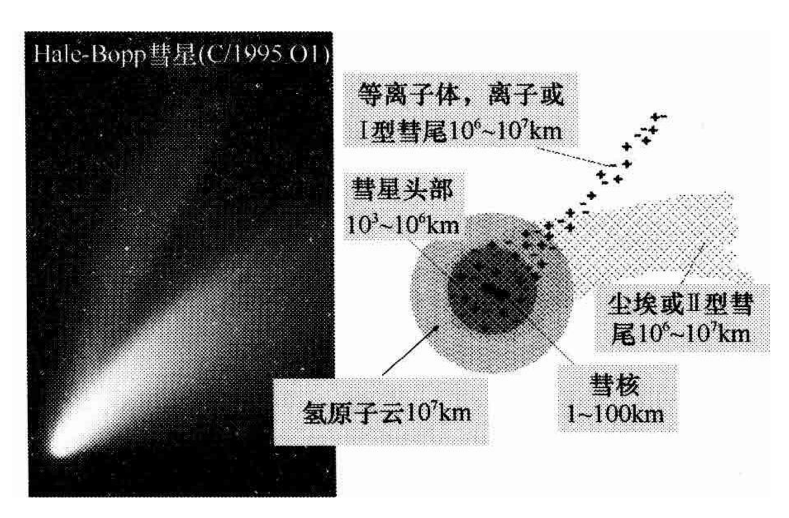


图 5.1　活动彗星的结构。太阳大约在与离子彗尾相反的方向。

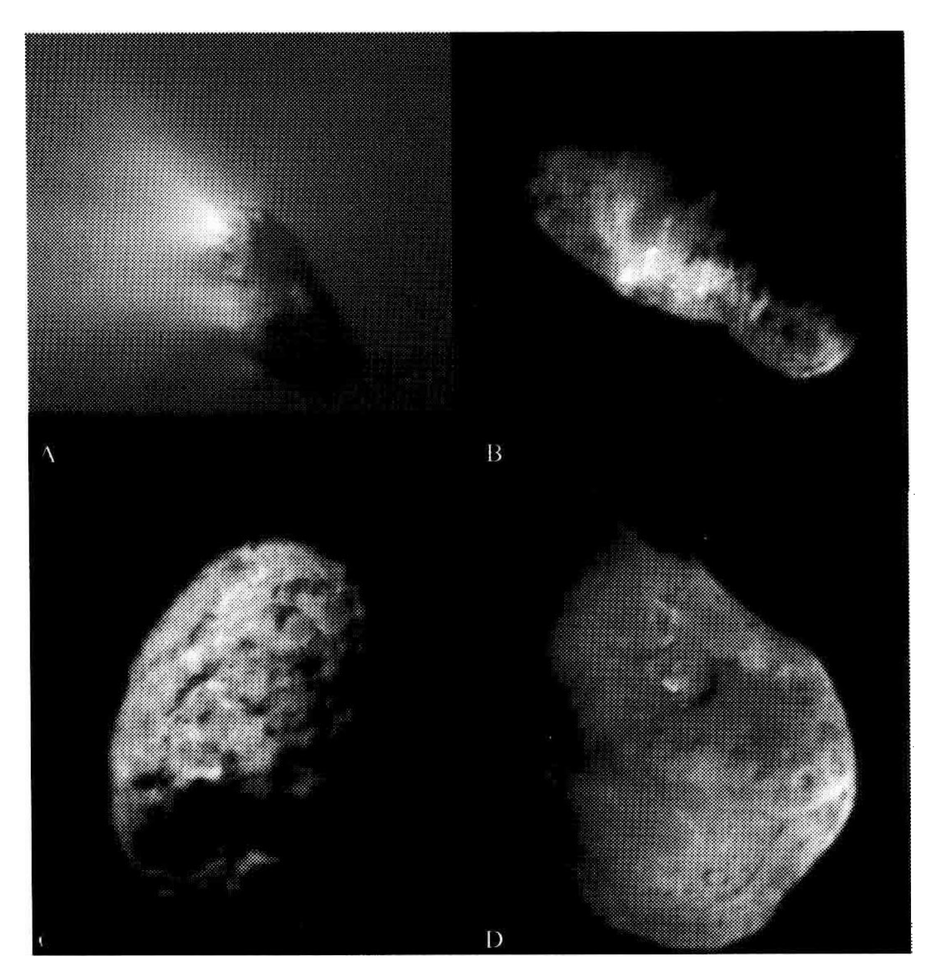


图 5.2　2007 年及以前发现的 4 颗彗星的核的照片。A. Halley 彗星的核，如 1986 年 ESA 探测器 Giotto 于几千千米的距离处所见（ESA 和 MPAe 提供）。B. P/Borrelly 彗星的核，2001 年由 NASA 探测器 Deep Space1 于 3400 km 处观测到（NASA 提供）。C. P/Wild 彗星的核，由 Stardust 在 236 km 处观测到，它是 NASA 探测器在近天体探测飞行时带回的收集的尘埃样本（NASA 提供）。D. P/Temple 1 彗星的核，由 NASA 探测器 Deep Impact 于 3000 km 处观测到（NASA 提供）。

5.3　彗星的化学组成

5.3.1　遥感

通过大型望远镜，得到彗星头部的无线电及红外图谱，证实 H_2O 是彗星冰的主要成分，但也探测出其他约 30 种含量不那么多的分子：含量丰富的 CO（为水的 1%～30%）和 CO_2（5%）及许多种与前生物化学反应相关的分子——HCN、NH_3、H_2CO、H_2S（约百分之几），HC_3N、CH_3CN、NH_2CHO、CH_3CHO、H_2CS…（0.01%～1%）。其中许多种物质在水溶液中都具有重要意义：HCN 是合成腺嘌呤及其他核酸碱基的关键分子，H_2CO 用于合成糖类（甲醛聚糖反应）、NH_3、HCN、H_2CO 及其他醛类，使得氨基酸和相关物质的

Strecker 反应能够进行（伴随有 CO_2 的 Buchere-Berg 合成）。由对彗星的观察及彗星之间的不同，表 5.1 列出了几种典型的挥发性分子的相对含量及不同彗星之间的差异[11]。

表 5.1 彗星上探测到的分子及一些上限

彗星挥发物种类	分子	海尔·波普(Hale-Boop)中含量	彗星间含量变化	已知彗星数+上限	
H	H_2O	100			水
	H_2O_2	<0.03			过氧化氢
C,O	CO	23	<1.4-23	9+8	一氧化碳
	CO_2	6	2.5-12	4	二氧化碳
C,H	CH_4	1.5	0.14-1.4	8	甲烷
	C_2H_6	0.6	0.1-0.7	8	乙烷
	C_2H_2	0.2	<0.1-0.5	5	乙炔
	C_4H_2	0.05?			丁二炔
	CH_3C_2H	<0.045			丙炔
C,O,H	CH_3OH	2.4	<0.9-6.2	25+2	甲醇
	H_2CO	1.1	0.13-1.3	18+3	甲醛
	CH_2OHCH_2OH	0.25			乙二醇
	HCOOH	0.09	<0.05-0.09	3+2	甲酸
	$HCOOCH_3$	0.08			甲酸甲酯
	CH_3CHO	0.025			乙醛
C,O,H 上限	H_2CCO	<0.032			乙烯酮
	$c\text{-}C_2H_4O$	<0.20			环氧乙烷
	C_2H_5OH	<0.1			乙醇
	CH_2OHCHO	<0.04			乙醇醛
	CH_3OCH_3	<0.45			二甲醚
	CH_3COOH	<0.06			乙酸
N	NH_3	0.7	<0.2-1	4	氨水
	HCN	0.25	0.08-0.25	32+0	氢氰酸
	HNCO	0.1	0.02-0.1	4+2	异氰酸
	HNC	0.04	<0.003-0.035	12+3	异氰化氢
	CH_3CN	0.02	0.013-0.035	9+2	甲基氰
	HC_3N	0.02	<0.003-0.03	3+7	丙炔氰
	NH_2CHO	0.015			甲酰胺
N 上限	NH_2OH	<0.25			羟胺
	HCNO	<0.0016			雷酸
	CH_2NH	<0.032			甲亚胺

续表

彗星挥发物种类	分子	海尔·波普(Hale-Boop)中含量	彗星间含量变化	已知彗星数十上限	
	NH_2CN	<0.004			氨基氰
	N_2O	<0.23			氧化亚氮
	NH_2-CH_2-COOH	<0.15			甘氨酸
	C_2H_5CN	<0.01			氰乙烷
	HC_5N	<0.003			丁二炔氰
S	H_2S	1.5	0.13-1.5	15+5	硫化氢
	OCS	0.4	<0.09-0.4	2+5	氧硫化碳
	SO	0.3	<0.05-0.3	4+1	一氧化硫
	SO_2	0.2			二氧化硫
	CS_2	0.17	0.05-0.17	15+3	二硫化碳
	H_2CS	0.02			硫甲醛
	S_2	0.005	0.001-0.005	5	二硫
	CH_3CH	<0.05			硫代甲醇
	NS	0.02	<0.02-0.02	1+1	硫化氮
P	PH_3	<0.16			磷化氢
金属	NaOH	<0.0003			氢氧化钠
	NaCl	<0.0008			氯化钠

提到彗星的矿物组成，远距离红外光谱及 ISO 卫星显示了镁橄榄石的存在，这是一种富含镁的结晶硅酸盐[12]。结晶和非结晶的硅酸盐同时存在。最近对这些数据进行了再分析，尝试对碳进行检测，对从中是否能探测到多环芳烃（PAH）仍存在争议。

最近，Deep Impact 号提供了一组有关彗星的新的原始数据。2005 年 7 月 4 日，一个 370kg 的撞击物以约 10km/s 的相对速度与 9P/Temple 1 彗星相撞。撞击物和携带它的宇宙飞船都拍下了彗核的照片，虽然多数的科学测量都是通过地球上的望远镜进行的。这次撞击很成功，带来了有关彗核性质的大量信息，但没有探测到新的挥发性有机组成。在这次撞击期间，彗星的活动在很多方面都与自然的彗星爆发相似。

5.3.2　就地观察

1986 年，Giotto 及 Vega 航天器经过哈雷彗星（Halley）时，其上装载的质谱仪提供了彗星尘埃粒子组分的直接信息。它们分析了耐热的有机粒子，之后将其命名为 C、H、O、N 粒子（来自碳、氢、氧、氮原子）和硅酸盐结晶。质谱仪显示了高分子质量的分子的存在，可能包括 H_2CO 的多聚体[15]。仪器检测到了

大量这类高分子质量物质，但由于质谱仪分辨率较低，许多不同种的物质被混在同一测量通道内。多数情况下单独鉴定是不可能的[16,17]。

彗星内尘埃和冰的含量是在同一数量级的，相差不超过 10 倍，彗星之间的不同可能反映了原始组分的不同和（或）由于在太阳系内轨道不同而引起的演化。在尘埃粒子中，硅酸盐和耐热有机物量差不多。

5.3.3 样本返回

星尘（Stardust）号是美国宇航局最早以收集彗星样本（81P/Wild 2）并将其带回地球为目标的航天器[18]。在太阳系内这是一颗比较新的彗星，因为它只有几次飞近了太阳。当彗星飞过一种密度非常低的硅基气凝胶时，它的尘埃粒子就被捕获了，这种气凝胶的设计能使捕获的物质减速，以使加热或热解对有机分子可能造成的任何改变降到最小。一旦被带回地球，马上就对这些粒子中的一部分进行分析，使用最先进和最敏感的仪器，分析其物理性质和化学组成。这些物质的其余部分留待未来用新技术进行分析。

81P/Wild 2 彗星超过 10 000 个粒子（1～300μm）被捕获并被带回地球，总质量约为 3×10^{-4} g[19]。最初的分析显示，这些彗星物质具有极复杂的组成。

Stardust 样本包括了结晶和非结晶的硅酸盐，如橄榄石和辉石。结晶矿物质的存在表明，其中一些物质是在非常高的温度下形成的（高于 1000K），而一部分硅酸盐处于非结晶态，说明它们经历了完全不同的热力学过程。这些结果表明，彗星物质由来源不同的物质混合而成，在太阳星云中混合得相对较好。

从有机化学的角度看，对这些粒子的最初分析探测到了有机组分，与含碳陨星相比，它含更多的氧和氮。这些样本非常不均匀，其 N/C 的范围为 5×10^{-3}～1。检测出了芳香族分子，如臭樟脑（$C_{10}H_8$）、菲（$C_{14}H_{10}$）和芘（$C_{16}H_{10}$），但样本中芳香族的含量比陨石和行星间的尘埃粒子要低。红外光谱鉴定出了 CH_2、CH_3，芳香族 CH、OH 和 C=O 基团。XANES（X 射线吸收近边结构谱）分析也检出了羰基、腈和氨基酸功能基团[20,21]，并对甲胺（CH_3-NH_2）、乙胺（$CH_3CH_2NH_2$）、甘氨酸（$NH_2CH_2CO_2H$）进行了特异性检测，但结果仍不成熟，除非进行更多的特异性同位素测量。若想得到对 Stardust 样本的最终结论，仍需数年极为认真的努力。

5.3.4 实验室工作

为了对更复杂且挥发性更小的物质有更好的理解，我们也可以转向实验室工作。这样的实验要遵循以下原则：通过对彗星头部及星际间介质中含量最丰富的物质的观察，可以推断出彗核中冰的可能组成。关键物质的气态样本可以在真空中沉积于一个较冷的层面，并在被 UV 光子或带电粒子沉积之时或之后进行扩

散。凝缩的冰有时候仅仅变暖，而不扩散。将样品加热进行分析时，留在层面上的耐热有机残留就是挥发性的升华物（图 5.3）。Mayo Greenberg 设想了这类实验，并把这种残留称为“黄色物质”[22]。

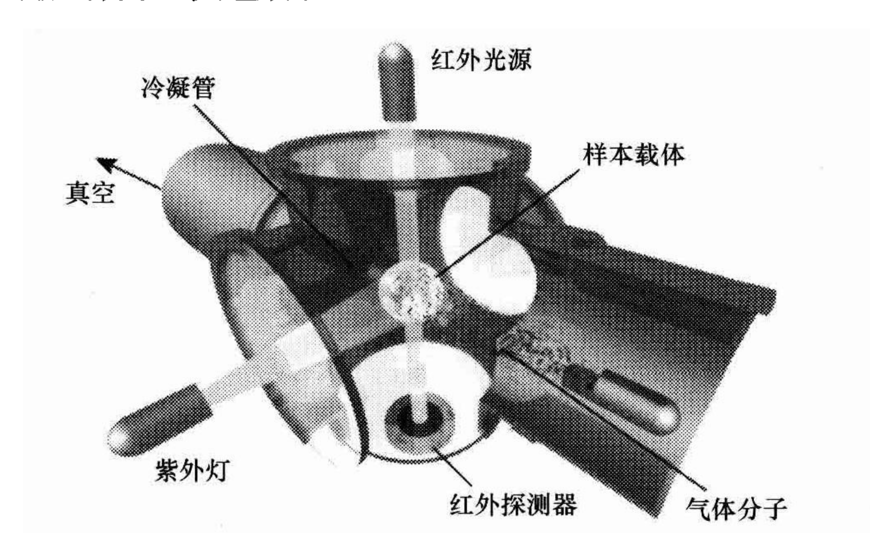


图 5.3　一个典型的实验装置，能够光解彗星冰类似物，在一个低温恒温器（紫外灯可以用离子或电子枪替换）中，将彗星冰类似物由气体混合物在一个冰冷的样本载体上冷却到 10 K 得到。可以经红外质谱直接分析冰的演化过程。室温下的残渣可以收集起来用于进一步的分析，如 GC-MS（气相色谱-质谱联用）、HPLC（高效液相色谱）和其他许多方法。图片由 Jan Hendrik Bredehoft，百里门大学，thoralf@uni-bremen.de 提供。

在这类实验室模拟中能产生多种多样的有机物，这点是值得注意的，但对它们的鉴别常常不完全[23]。复杂分子的性质由冰的组成和能源性质决定。在实验中使用的三种能量过程（热循环、UV 光解、能量粒子辐射）可能对冰混合物进行作用，要么作用于星际星云中被冰包被的尘埃粒子，要么作用于太阳星云里正在形成的冰冻小行星体，要么作用于太阳系内彗星中冰的外层。怎样限制不同过程对宇宙中冰的影响程度是一个很费解的难题。根据能源不同而造成的产物不同，可以给出彗星物质及彗星以前的信息。对这个问题仍在进行研究。

从太空生物学角度来看，值得一提的是，在对冰混合物进行 UV 照射后，获得的残留中有大量氨基酸（如甘氨酸、丙氨酸、肌氨酸、缬氨酸、脯氨酸、丝氨酸等），这些冰混合物包括 H_2O：NH_3：CH_3OH：HCN 和 H_2O：CH_3OH：NH_3：CO：CO_2。未被水解的残留物（在分析试剂中没有加入液态水）只产生了微量的甘氨酸。对其他氨基酸的检测需要在非常强烈的条件（HCl$\geqslant$6 mol/L 且 $T\geqslant$100℃）下进行酸解[24,25]。因此目前未能确定：①氨基酸是否真的存在于实验室残留物进而亦存在于彗星冰中，还是“只”合成了氨基酸前体；②在彗星撞击早期地球的原始海洋的过程中，氨基酸前体是否可以被酸解变成真正的氨基酸。

通过实验室模拟，也探究了手性问题。有报道说，对固体亮氨酸的外消旋混

合物进行不均匀的真空 UV 光解，可导致这种氨基酸某种形式的对映体过量[26]。但是，若在冰混合物中，直接用圆偏振光合成氨基酸，却没有发现这样的对映体过量[27]。对这个问题进一步的研究，已在 2007 年法国同步加速器 SOLEIL 开放后进行（DESIRS 光路）。

遥感和就地观察只能探测彗星的气体，样本返回可研究有限量的彗星物质，实验室工作可以模拟彗星冰，具有雄心壮志的下一步计划是登陆彗星，研究其组成，这将由欧洲航天局（ESA）的 Rosetta 号完成。

5.4 Rosetta 2014——实验室与彗星的约会

Rosetta 号由两部分组成：一个绕核运转的卫星和一个称为 Philae 的着陆器。它发射于 2004 年 3 月 2 号，将于 2014 年 8 月到达目标——67P/Churuymov-Gerasimenko 彗星。Philae 预计于 2014 年 11 月在彗核上着陆。卫星上的机载仪器将能够对从彗核释放的挥发性及耐热的组分进行前所未有的分析。Rosetta 上装载的两台质谱仪（ROSINA：Rosetta 卫星离子和中子分析质谱仪；COSIMA：彗星二次离子质谱仪），将收集和分析仅距彗核表面 1 km 处的气体和尘埃，希望能够达到足够近，以研究几乎没有产生变化的彗核物质。在一个广阔的电磁辐射光谱范围内，4 台光谱仪也将探测彗星环境的组成（OSIRIS：光学分光及红外遥感成像系统；ALICE：紫外光谱仪；VIRTIS：视觉和红外热成像光谱仪；MIRO：Rosetta 卫星微波仪）。在 Philae 上，太空生物学方面最丰富的信息将来自 COSAC（彗星采样及组成研究），为气相色谱质谱联用仪。在这一章，我们强调仪器 COSAC、COSIMA、VIRTIS 和 MIRO，因为多数新的有机分子的发现都会归功于它们。

5.4.1 COSAC

COSAC 是一台气相色谱（GC）-质谱（MS，在这里为线性飞行时间质谱仪）联用仪。它有 8 个色谱柱，每个都与单独的探测器相连（TCD），但也可以将它们与质谱仪相连[28]。之前获得的结果要归功于 Puma、Giotto 和 Stardust 航天器直接进行的质谱测量，它们“仅仅”给出了所有分子同时混合时的质谱。COSAC 将用色谱进行初步分离，这将在一定程度上确定组分，因为组分的鉴别既要根据它们单独的质谱，也要根据其保留时间。在钻探彗星表面后将会收集样本，在进入分析系统之前，要将它们加热到不同温度。为了降解最耐热的组分，可能要进行热解（高于 600℃），来对碎片进行气相分析。这 8 个色谱柱中，有 3 个专门用于分析手性分子，以区分对应异构物。挑选出另外 5 个色谱柱，以便检测最大量的分子。另外，几个色谱柱同时对单个样本进行分析，能够通过比较对数据进行更好的分析，从而提高对数据解释的可信度。

许多分子都被选择进入色谱柱（已在彗星气体中检测到的分子、星际间介质中的分子或来自实验室对彗星冰模拟时产生的分子）。必须提到的是，氨基酸及其他高质量组分（如草酸、尿素等）都是不能被检测到的，因为它们的挥发性不够，不能在气相中进行分析。对这类组分的 GC 分析需事先对其进行“衍生作用”（能使目标物质挥发性更强的化学过程）。这个步骤对 COSAC 仪不适用，但是在未来的火星探测实验中，让衍生作用能够进行的研究已经开始。COSAC 对彗核的分析将由 CIVA（红外分析）和 MODULUS（同位素测量）完成。

5.4.2 COSIMA

COSIMA 是一台时间飞行二次离子质谱仪（TOF-SIMS），可对彗星尘埃粒子进行就地分析[25]。它将以金属彗星粒子为目标（银、金、钯和铂）进行收集，一个光学系统（COSISCOPE）将对这些粒子进行定位，再由 TOF-SIMS 进行更多分析。COSIMA 的质量分辨能力（$M/\Delta M$）约为 2000，分析粒子的分辨率约为 50μm。TOF-SIMS 技术对彗星表面组成非常敏感，只分析样本的第一个单分子层。TOF-SIMS 谱图非常难以解读，因为它们包含了非常大量的信息，可同时显示质量达 1000 amu 原子质量单位（Atomic Mass Unit）的元素和分子。

不同于 COSAC 上的分析，COSIMA 质谱仪在分析之前不将不同分子进行分离，因此会出现极复杂的图谱。但很容易对有机和矿物组分进行可靠的区分（富含 H 的分子，如有机物，通常轻微高于单位数，$m=n+\delta$，$n\in N$；而矿物质质量通常轻微低于单位数，$m=n-\delta$）。任何特异性的鉴别可能都是冒险的，除非用地面仪器进行精细的校准，这一点目前已在进行。这台仪器有非常高的质量和空间分辨率，可对相同粒子的不同点进行组分分析。另外，COSIMA 可以分析不可挥发的组分，它们是无法被 COSAC 探测到的。

由 COSIMA 进行的分析将由 ROSINA 的气象分析完成[30]。这台基于质谱测量的仪器将通过两个传感器确定彗星大气和电离层的组成。ROSINA 的质量测量范围很宽（1～300 amu），且具有高质量分辨率，$M/\Delta M>3000$，可将 N_2 中的 CO 和 ^{12}CH 中的 ^{13}C 分辨出来。

5.4.3 VIRTIS 和 MIRO

两台光谱仪 VIRTIS 和 MIRO 将从卫星上观测彗核及彗星头部，它们将提供彗星物理条件和物理过程，以及彗星化学组成的相关信息。

VIRTIS 是一台 UV/可见及红外光谱仪[31]。两个通路分别覆盖波长为 0.2～1μm 和 1～5μm 的范围。有关彗星表面性质，以及彗星头部释放的分子（“母体分子”），在它们被光解、电离或与其他物质反应之前，红外通路将提供一些相关信息。

对于彗核，将以高空间分辨率远距离定位其上的矿物质和冰。预期会有硅酸盐和氢氧化物；如果存在，层状硅酸盐（“黏土”）将通过它的水和 OH 吸收带显出踪迹。水冰和其他的冰（如 NH_3、CO_2 或 H_2S）可由其红外光谱特性而鉴别。碳氢化合物（CH_4、C_2H_2、C_2H_4、C_2H_6、C_3H_8）形成的冰也可由其光谱范围确定。这些简单的碳氢化合物中，有一些能够聚合。这样的含碳物质将使彗星光谱发生红移，更精确的鉴定则需要更高的信噪比（>100）。

在彗星头部，除了粒子产生的固体特性，更精细的特性来自于其中的气体。一个主要任务是，通过高光谱分辨率（$\lambda/\Delta\lambda>2000$），鉴定在 3～4$\mu$m 范围内泄出的碳氢化合物。其红外光谱也含丰富的重要母体分子——H_2O、CO_2、H_2CO、CH_3OH、CO 的旋转与摆动放射。氘化水 HDO 也可经长时间检测被鉴定出来；HDO/H_2O 是衡量彗星对构建地球海洋重要性的关键信息。目前只有三颗彗星（Halley、Hyakutake 和 Hale Bopp）的数据可用，其水中的 D/H（约 3×10^{-4}），为 SMOW（标准平均海水，约 1.5×10^{-4}）的两倍。但是，我们认为仅仅对这几颗彗星的观察不能代表所有彗星，而且在这几颗彗星的不同位置，该比例是否相同也是个问题。

VIRTIS 的紫外线通路主要对彗星头部反应的产物（“子分子”）敏感：分子基团（OH、CN、C_2、C_3、NH、CH）、离子（CO^+、CH^+、H_2O^+、N_2^+）和小分子（CO）。一些产物可保留母体的性质，如果认为这些物质是彗星粒子中可能存在的复杂有机物的光解或热解产物的话，这一点就更值得注意了，因此，可以直接探测到多聚体的存在，如 H_2CO 的多聚体聚甲醛（POM）或 HCN 多聚体。

MIRO 由两个无线电接收器组成，工作波长分别在毫米级和亚毫米级（1.6mm和 0.5mm，或 190 GHz 和 562 GHz）[32]。由于天线尺寸较小（30cm），且接收器没有冷却设备，将只能探测到主要的彗星物质 H_2O（及其同位素 $H_2^{17}O$ 和 $H_2^{18}O$）、CO、CH_3OH 和 NH_3。彗核排出气体的精确物理数据，以及由这些数据推论出的彗星头部发育状况，将用于这些数据的详细建模，建模过程将通过更巨型的地面仪器进行，如西班牙的 IRAM 30m，墨西哥的 LMT 50m，Virginia 的 GBT 110m，或建于智利地下的 ALMA 干涉仪（直径 12m，64 根天线）。NH_3 自身是很难从地面观测到的，而且通过厘米级无线电波或近红外线，也只在很少几颗彗星上探测到了。在其地球轨道上，ODIN 卫星试着用与 MIRO 相同的 572GHz 对两颗彗星上的 NH_3 进行了探测。太阳系的冰中，NH_3 的存在降低了水冰的熔点，因而在比纯水熔点低的温度下，仍能有液态水的存在。

5.5 彗星与生命

5.5.1 前生物分子的传送

前生物分子的外界来源，一方面来自大物体、彗星和小行星的撞击；另一方面来自彗星或小行星产生的大小为微米到毫米的行星间尘埃粒子的缓慢着陆。这两种机制的传递条件非常不同，且对其结果非常难以估计，但作为前生物物质补充原料，它们可能都具有数量上的重要性[33]。在这里我们主要讨论彗星。

A. 由撞击引起的强烈传送

尽管彗核对地球的撞击是高能的过程，似乎所有的分子都应该毁灭，但是实验室实验和数字模拟显示，一次撞击可能给早期地球带来一些氨基酸（如天冬氨酸和谷氨酸），其含量与在富含 CO_2 的大气中 Miller-Urey 合成的量类似[34]。如果考虑到小天体，如火星、木卫二 和月球，尽管其较低的撞击速率有利于分子的保留，但由于其较小的重力，较多的撞击碎片可能逃逸，导致产量更低；这种生产过程对火星来说还是相对有效的，但对更小的木卫二和月球则不然。然而，尽管木卫二的形成历史不允许它含有或就地形成有机物，木卫二仍可通过这样的彗星撞击获得一些有机分子。

应该提到，小行星（有些富含有机物）没有类似的有机物传送效率，因为它们更坚硬，孔隙更少，且内部密度更低。撞击时可能达到更高的温度，导致复杂有机物的毁灭比例更高。

B. 缓慢减速的同温层星际尘埃粒子（SIDP）和软着陆的微陨石

小粒子（10μm）可被大气层有效减速，然后“四处飘浮”足够长的时间，使空中收集器能够收集它们。这些绒毛似的粒子被称为 Brownlee 粒子，其来源可能是彗星。有时候大一些的 100～200μm 的粒子到达地面，它们被称为微陨石。可以在极地的冰中找到它们，在那里它们更容易被发现，且较少被地球粒子污染。SIDP 和微陨石是来到地球的地外物质的通常来源；估计一年有 10 000～40 000t，它们的重要性仅次于大撞击（若将其贡献平均到地质年代）。

这些粒子的碎片富含碳，甚至比含碳的球粒状陨石还高。在一个富含碳的微陨石样本中，发现了两种氨基酸——AIB 和异缬氨酸，后者不存在于地球上的活生命体。

C. 大气电击合成

除了它们携带并可能传递到地球的有机分子，下落的天体还能将能量带入大气，这会引起电击，导致化学合成发生，这比还原性的 Urey-Miller 大气效率更高，但不如弱还原性的大气（$H_2/CO_2<0.1$）。

5.5.2 彗星是生命破坏者?

根据所处时代（及不同模型），在千米大小的彗星中，可能有1%～50%撞击了地球。通过传送或电击引导合成有机分子，这样的撞击可能有利于生命的出现，但是，它们对初期生命也可能是有害的甚至致命的。两个可能具有彗星性质的事件仍在争论中：①1908年的Tunguska事件（一个数十米的天体），摧毁了2000km^2的西伯利亚森林；②65Myr前的K/T事件，造成了Chixculub陨石坑（来自于一个大于10km的天体），也是大规模灭绝生物（包括恐龙）的最受欢迎的解释[35]。

5.5.3 彗星上有液态水?

彗星上本身能够进行先进的前生物化学反应而产生生命吗？这个问题是有意义的，尤其是如果彗星上存在液态水的话。液态水需要足够高的温度和压力。在彗星形成时，增长过程主要加热了彗星表面。另外两种机制可以在彗核内部产生热量——放射加热和非结晶到结晶的冰的不可逆转化。在太阳系早期，铝显然是最重要的放射加热元素（其重要性约为其他元素的500倍），其半衰期为700 000年，在星际介质中广泛分布，可能在原始太阳星云中存在。如果彗星足够大，可以避免迅速降温，它就可以通过热传导和气体扩散加热彗星内部；但只有当彗星形成得非常迅速（小于几个半衰期）时，这种情况才可能发生。根据数字模型，10km大小的彗星可以含液态水达10^5年[36]。

5.6 结论

彗星可将元素（如碳等）带到地球，同时带来的还有水（量仍不确定），此外，它们含有简单的有机分子，可能也含有更复杂的分子，因此，通过释放尘埃粒子（有些在进入地球大气时毁灭了）、大气化学合成和彗核撞击地球时可能发生的直接传递，彗星可能有助于早期地球对前生物分子的富集。在局部或全球尺度上，这些撞击也可能强烈影响后来生命的发展。

目前的争论关联了很多问题，我们希望在未来的这些年里找到答案。空间任务可以为这些问题带来重要的结果。

- 彗星物质及后来的彗星，是什么时候、在哪里形成的?
- 彗星物质和星际间物质有多相似?
- 彗星含氨基酸和其他生命的基本组分吗?
- 这些分子是呈对映体过量，还是外消旋混合物?
- 彗星对地球的贡献（海洋中的水、碳和前生物分子）到底是多少?
- 彗星对生命及其出现和发展的全球性影响到底是什么?

• 奥尔特云或 Kuiper 带的存在，是我们太阳系中的特例，还是在太阳系外主序星中的普遍存在？

考虑到彗星可能在地球生命起源和进化中起到的关键作用，我们有必要知道，一颗可能具有可居住行星的恒星周围，是否也环绕着彗星。新的研究课题将更多关注这些系外彗星（轨道围绕其他恒星的彗星）。最近发现一颗邻近恒星 β Pictoris 是有圆盘环绕的主序星之一，其光谱特性可以被解释为众多“蒸发体”的持续冲击。通过最近野心勃勃的太空实验如 Corot 项目，我们可以期待更多的信息，它的任务是，在许多其他恒星周围，探测地球大小的行星，或更直接探测彗星[37]。

参考文献

扩展阅读

1. Crovisier J, Encrenaz T. Comet Science: The Study of Remnants From the Birth of the Solar System. Cambridge: Cambridge University Press, 2000.
   • 覆盖面广、插图丰富的浅易的彗星介绍。
2. Despois D, Cottin H. Comets. Potential sources of prebiotic molecules for the early Earth. In: Gargaud M, Barbier B, Martin H, Reisse J, eds. Lectures in Astrobiology. Berlin: Springer Verlag 2005; 1:289-352.
   • 比本章详细且参考文献较多。
3. Irvine WM, Schloerb FP, Crovisier J et al. Comets: A link between interstellar and nebular chemistry. In: Manning V, Boss B, Russel S, eds. Protostar and Planets IV. Tuscon: University of Arizona Press 2000a:1159.
   • 含详细彗星化学。

Two recent multi-author books are highly recommended for detailed research work.

4. Thomas PJ, Hicks RD, Chyba CF et al. Comets and the Origin and Evolution of Life. 2nd edition. New York: Springer 2006.
   • 详细讨论彗星撞击及水的问题。
5. Festou MC, Keller HU, Weaver HA. Comets II. Tuscon: University of Arizona Press 2004.
   • Reviews much of cometary science beyond the astrobiological interest.
6. Newburn RL, Neugebauer M, Rahe JH. Comets the Post-Halley Era. New York: Springer 1991.
   • The two volumes of Comets in the Post-Halley Era, despite their age, are still a very good detailed introduction to cometary science.

专题文献

7. Whipple FL. A comet model I. The acceleration of Comet Encke. Ap J 1950; 111:375-394.
8. Chamberlin TC, Chamberlin RT. Early Terrestrial conditions that may have favored organic synthesis. Science 1908; 28:897.
9. Oro J. Comets and the formation of biochemical compounds on the primitive Earth. Nature 1961; 190:389-390.
10. Festou MC, Keller HU, Weaver HA. Comets II. Tuscon: University of Arizona Press, 2004.
11. Bockelée-Morvan D, Crovisier J, Mumma MJ et al. The composition of cometary volatiles. In: Festou M, Keller HU, Weaver HA, eds. Comets II. Tucson: University of Arizona Press 2004:391-423.
12. Crovisier J, Leech K, Bockelée-Morvan D et al. The spectrum of Comet Hale-Bopp (C/1995 O1) observed with the Infrared Space Observatory at 2.9 Astronomical Units from the Sun. Science 1997; 275:1904-1907.
13. Lisse CM, Kraemer KE, Nuth III JA et al. Comparison of the composition of the Tempel 1 ejecta to the dust in Comet C/Hale-Bopp 1995 O1 and YSO HD 100546. Icarus 2007; 187:69-86.
14. Lara LM, Boehnhardt H, Gredel R et al. Behavior of Comet 9P/Tempel 1 around the Deep Impact event. A and A 2007; 465:1061-1067.
15. Huebner WF. First polymer in space identified in comet Halley. Science 1987; 237:628-630.
16. Kissel J, Krueger FR. The organic component in dust from comet Halley as mesured by the PUMA mass spectrometer on board Vega 1. Nature 1987; 326:755-760.
17. Mitchell DL, Lin RP, Carlson CW et al. The origin of complex organic ions in the coma of comet Halley. Icarus 1992; 98:125-133.
18. Brownlee D, Tsou P, Aleon J et al. Comet 81P/Wild 2 under a microscope. Science 2006; 314:1711-1716.
19. Horz F, Bastien R, Borg J et al. Impact features on Stardust: Implications for comet 81P/Wild 2 dust. Science 2006; 314:1716-1719.
20. Keller LP, Bajt S, Baratta GA et al. Infrared spectroscopy of comet 81P/Wild 2 samples returned by Stardust. Science 2006; 314:1728-1731.
21. Sandford SA, Aleon J, Alexander CMO et al. Organics captured from comet 81P/Wild 2 by the Stardust spacecraft. Science 2006; 314:1720-1724.
22. Greenberg JM. What are comets made of? A model based on interstellar dust. In: Wilkening LL, ed. Comets. Tucson: University of Arizona Press 1982:131-163.
23. Colangeli L, Brucato JR, Bar-Nun A et al. Laboratory experiments on cometary materials. In: Festou MC, Keller HU, Weaver HA, eds. Comets II. Tuscon: University of Arizona Press, 2004:695-717.
24. Bernstein MP, Dworkin JP, Sandford SA et al. Racemic amino acids from the ultraviolet photolysis of interstellar ice analogues. Nature 2002; 416:401-403.
25. Muñoz Caro GM, Meierhenrich UJ, Schutte WA et al. Amino acids from ultraviolet irradiation of interstellar ice analogues. Nature 2002; 416:403-406.
26. Meierhenrich UJ, Nahon L, Alcaraz C et al. Asymmetric vacuum UV photolysis of the amino acid leucine in the solid state. Angewandte Chemie-International Edition 2005; 44:5630-5634.
27. Nuevo M, Meierhenrich UJ, Muñoz Caro GM et al. The effects of circularly polarized light on amino acid enantiomers produced by the UV irradiation of interstellar ice analogs. A and A 2006; 457:741-751.
28. Goesmann F, Rosenbauer H, Roll R et al. Cosac, The cometary sampling and composition experiment on Philae. Space Science Reviews 2007; 128:257-280.
29. Kissel J, Altwegg K, Clark BC et al. COSIMA, a high resolution time of flight spectrometer for secondary ion mass spectroscopy of cometary dust particles. Space Science Reviews 2007; 128:823-867.
30. Balsiger H, Altwegg K, Bochsler P et al. Rosina—Rosetta Orbiter spectrometer for ion and neutral analysis. Space Science Reviews 2007; 128:745-801.
31. Coradini A, Capaccioni F, Drossart P et al. Virtis: an imaging spectrometer for the Rosetta Mission. Space Science Reviews 2007; 128:529-559.
32. Gulkis S, Frerking M, Crovisier J et al. MIRO: microwave instrument for Rosetta Orbiter. Space Science Reviews 2007; 128:561-597.
33. Chyba CF, Hand KP. Comets and prebiotic organic molecules on early Earth. In: Thomas PJ, Hicks RD, Chyba CF, McKay CP, eds. Comets and the Origin and Evolution of Life. 2nd edition. New York: Springer, 2006:169-206.
34. Pierazzo E, Chyba CF. Impact delivery of prebiotic organic matter to planetary surfaces. In: Thomas PJ, Hicks RD, Chyba CF, McKay CP, eds. Comets and the Origin and Evolution of Life. 2nd edition. New York: Springer, 2006:137-168.
35. Zahnle K, Sleep NH. Impacts and the early evolution of life. In: Thomas PJ, Hicks RD, Chyba CF, McKay CP, eds. Comets and the Origin and Evolution of Life. 2nd edition. New York: Springer, 2006:207-252.
36. Podolak M, Prialnik D. The conditions for liquid water in cometary nuclei. In: Thomas PJ, Hicks RD, Chyba CF, McKay CP, eds. Comets and the Origin and Evolution of Life. 2nd edition. New York: Springer 2006:303-314.
37. Lecavalier Des Etangs A, Vidal-Madjar A, Ferlet R. Photometric stellar variation due to extra-solar comets. A and A 1999; 343:916-922.

6 陨星与生命起源前的化学

Sandra Pizzarello

6.1 陨星及其来源

陨星是地外物质的一部分，穿越了地球轨道，进入地球大气，并且到达地面，并没有被完全摧毁。已知其中一些是由火星、月球甚至水星在经历撞击后喷射出来的。但是对大多数来说，陨星是小行星体的碎片。它们是具有不同大小和形状的小行星（直径<1000km），遍布太阳系，但大部分集中在火星和木星之间（已识别了超过 2000 个）。根据提丢斯-波得定律（Titius-Bode law），行星应分布在离太阳的等距位置上，因而在小行星轨道上应当有一颗行星，所以有人讨论，小行星是否是一颗爆炸了的行星的残余，是否是被附近行星的重力作用拆裂的。根据现有的太阳系形成的模型，目前达成的共识是，它们代表了一颗没能形成的行星。当大行星木星在附近凝缩时，人们相信，它的重力作用影响了周围正在形成一颗行星的小天体，把其中的一些作为彗星抛到了更远的周边轨道，使小行星带上剩下的部分不能结合[1]。

小行星包括行星没能形成的幸运之处是，大量的小行星体没有经历过伴随着早期行星过程的地质变迁，比如不稳定物质的散失和岩石之间的互熔，从而保存了早期太阳系物质的原始记录。

但是，小行星带是一个拥挤的地区，经常出现碰撞，我们拍下的照片显示了这些小行星坑坑洼洼的表面（图 6.1），并且这些碰撞产生的块状物能够进入地球轨道，经常作为陨星到达地球。目前已经收集了超过 18 000 颗陨星；一小部

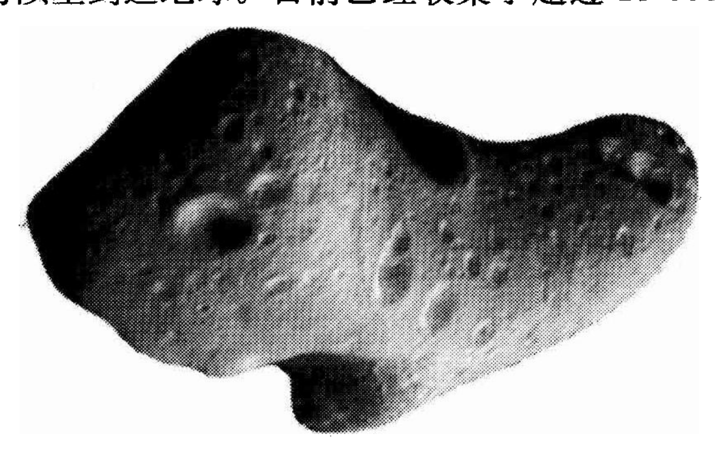


图 6.1　Eros 陨石，2001 年情人节由一个美国宇航局探测器发现。

分是下落时被记录的（记录它们的下落位置），其余的则是找到的，没有任何人目击到它们的下落。后面这一组里的多数都是最近的二十年内在南极洲发现的，日本和美国都在南极洲上拥有区域，并已为其搜寻派遣了科考队。

根据大体上的相似类型，陨星被分为几个不同的类和亚类，主要是铁陨石、石陨石和石铁陨石这几类。铁陨石为铁/镍合金块，来自那些足够大的小行星，它们从重力坍缩中得到了足够的热量，其物质根据密度不同而分离，得到了一个密度更高的铁核。这类陨星似乎普遍起源于小行星带距太阳更近的部分。石陨石外表为岩石，由不同种类的硅酸盐和其他矿物质组成，根据其矿物组成和地化改变程度而分类。

6.2　含碳的球粒状陨石及它们所含的有机物

含碳球粒状陨石（Carbonaceous Chondrites，CC）是石陨石的一个亚类，得名于它们中称为陨石球粒的圆形玻璃状内含物，以及其有机碳含量的不同百分比（1%～4%）。CC 具有原始的元素组成，这与太阳及整个宇宙类似，因为太阳是一颗相当一般的恒星。它们的外表为聚集的石头（图 6.2），即物质堆在一起，没有经过极端的温度或压力转化的痕迹。它们的主要组分是结晶得非常好的硅酸盐基质，内含有较大的矿物质，并有异质性的有机物散布其中[2]。

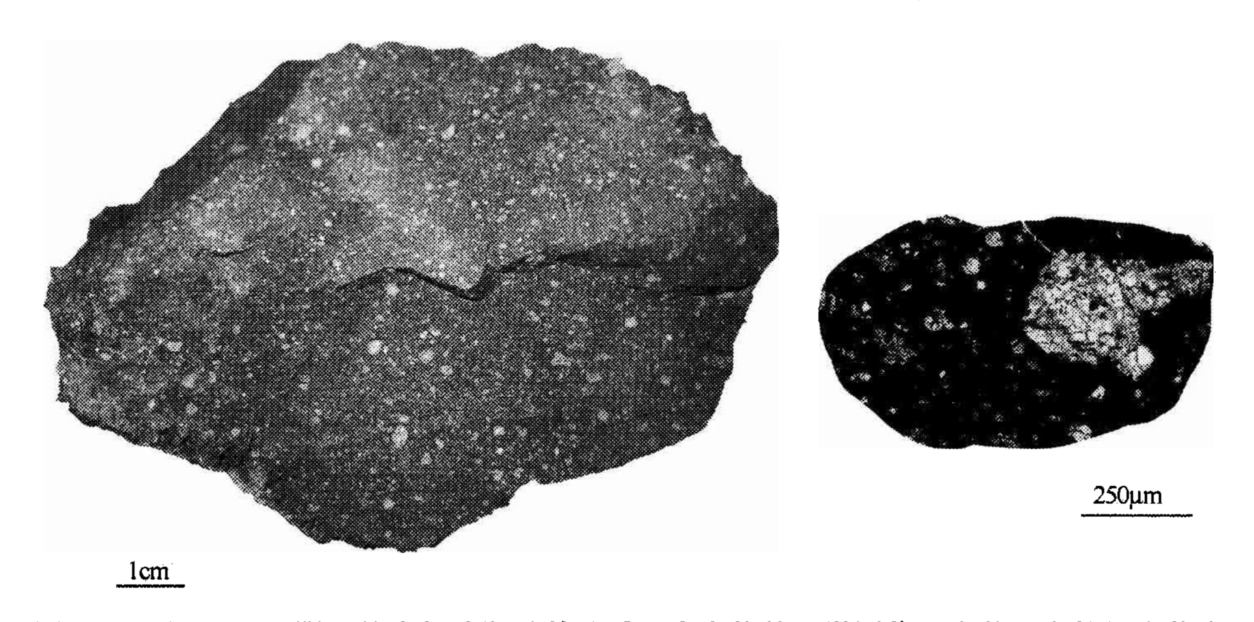


图 6.2　Murchison 陨石的内部片段及其基质和内含物的显微图像。由美国宇航局庄信太空中心 Dr. Michael E. Zolensky 提供（见彩插）。

由于 CC 特殊的原始性质及陨石的普遍分布，它们的有机物使我们能够直接分析无生命有机化学反应的天然样本，在生命出现之前不久，这些反应在太阳系中进行着。

表 6.1 给出了 Murchison 陨石的一般有机组成，它很大（超过 100kg），于

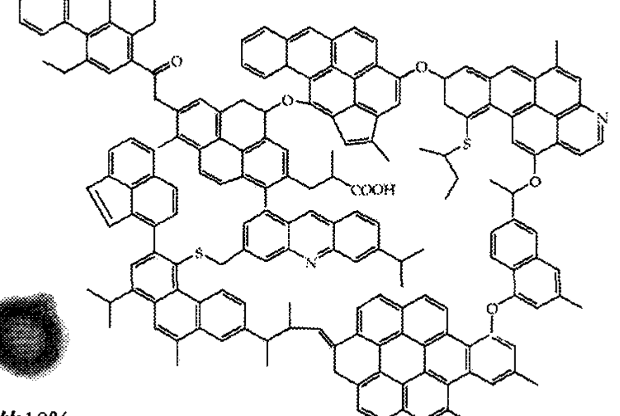
1969[a] 年落下，并从那时起就被广泛研究。有时形容陨石有机碳中较大的部分为油母岩类似物，因为，正与地球上的油母岩一样，这是一种不可溶的大分子物质，除C和H外，还由N、O和S组成，它的复杂构成还没有被详细了解。这些物质的大小主要由其光谱分析[b] 和分解研究推断，通过加热热解或氧化变成片段。人们确信它具有图 6.3 中显示的假定结构，具有芳香环的簇，由含S、N、O的脂肪链进行桥连，还带有周边分支和功能基团。这些不可溶的碳化合物是驳杂的，约 10%含有具不同元素组成的纳米结构、球状结构和纳米管；也在其中发现了微量的“外来”碳，它们可能形成于恒星的外部包层中（即在太阳系形成之前）。此外，最近发现这些物质不是完全不可溶的，并且可以在类似地球热液口处（300℃，100 MPa）的高温高压条件下释放单个分子。

表 6.1　Murchison 陨石的有机组分

碳酸盐矿物质（CO_3^{2-}）	2%～10%
不可溶有机碳	65%～80%
可溶有机成分	10%～20%

化学式：$C_{100}H_{71}N_3O_{12}S_2$

大分子物质：约80%：

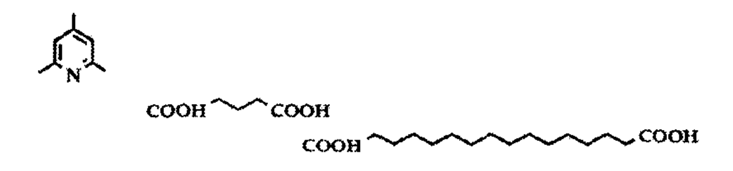


结构差异：约10%

例：纳米小球和小管

热液中含有的化合物：约10%

例：C_3-C_{17}二羧酸和甲基吡啶



来自星际微粒的外源碳：百万分之一

例：钻石，石墨，以及富勒烯；SiC;Si_3N_4

图 6.3　Murchison 陨石的不可溶有机碳组成。

其他许多与 Murchison 类型相同的含碳陨石，其有机碳的其余部分由可溶组

分构成，用水及溶剂抽提，可从陨石粉末中释放出来（表 6.2）[3]。就我们所知，正如从它们的名字和结构所见，许多单个组分的类型遍布生物圈，而且在陨石中也存在。

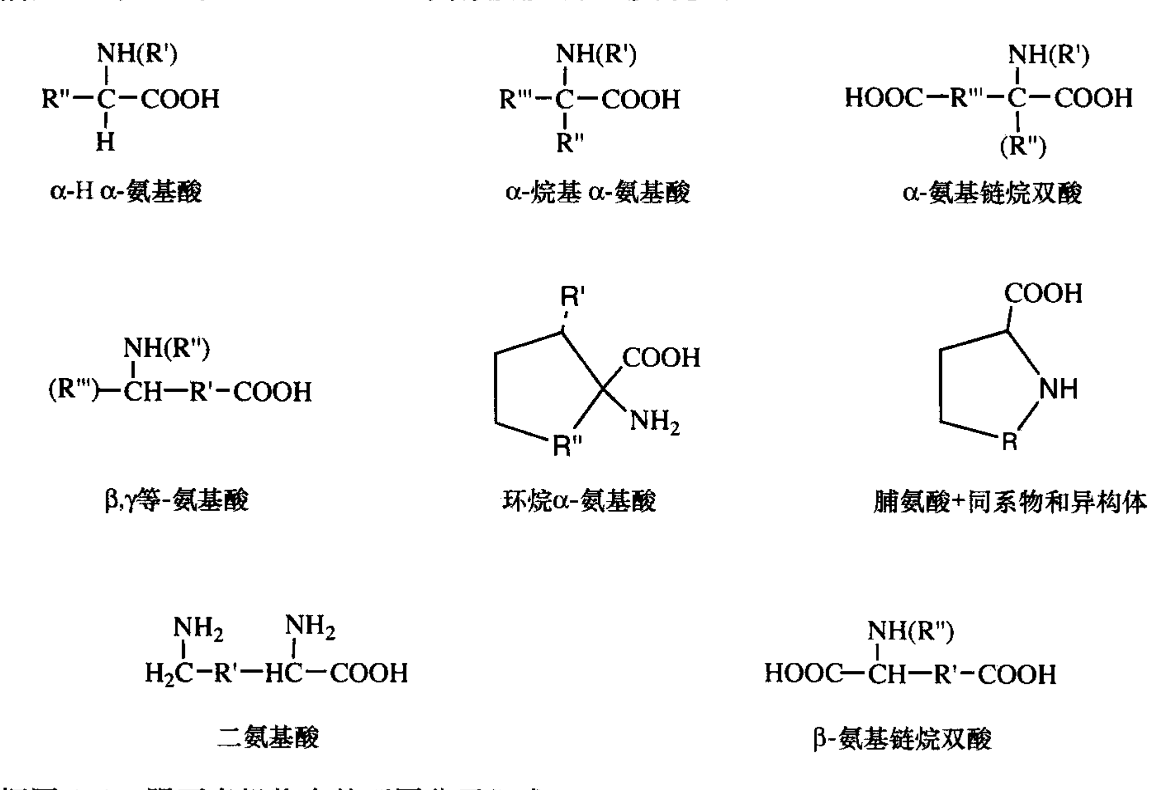
表 6.2　Murchison 陨石中的可溶有机化合物

化合物	典型结构	分子和典型分子
羧酸	$H_3C—COOH$	乙酸
氨基酸	NH_2 | $H_3C—CH—COOH$	丙氨酸
羟基酸	OH | $H_3C—CH—COOH$	乳酸
酮酸	O ‖ $H_3C—C—COOH$	丙酮酸
二羧酸	$OOC—(CH_2)_2—COOH$	琥珀酸
糖醇和糖酸	OHOH | | $H_2C—CH—COOH$	甘油酸
醇、醛和酮	$H_3C—CH_2OH$	乙醇
胺类和氨基化合物	$H_3C—CH_2NH_2$	乙胺
吡啶羟酸	COOH N	烟酸
嘌呤和嘧啶	NH_2 N N N N H	腺嘌呤
烃：脂肪族	$H_3C—CH_2—CH_3$	丙烷
芳香族		臭樟脑
极性	N	异喹啉

陨石的有机组成，即使量只有百万分之几 ppm（μg/g），也具有许多种不同形式，常常包含其化学基团的所有可能结构，并达到其可溶性的上限。

举例来说，地球上的所有蛋白质都由仅仅 20 种氨基酸组成，且全部是 α-氨基酸，然而在陨石上已经发现了超过 100 种氨基酸，包括了许多或全部含烃基链，一到两个羧基及一到两个氨基的可能化学结构（框图 6.1）。有 10 种氨基酸

在CC和地球蛋白质中都有发现：甘氨酸、丙氨酸、丝氨酸、苏氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、天冬氨酸和谷氨酸（丝氨酸和苏氨酸仅在最近来自庄信太空中心的GRA 95229号南极陨石上发现）。



框图 6.1 陨石有机物中的不同分子组成

其他在陨石和生物圈都有发现的有机物有烟碱酸（以烟碱形式存在，它是一大类氧化还原酶的辅酶）、一些单羧酸（如蚁酸、乙酸、丁酸，得名是由于它们分别是在蚂蚁、醋和沸腾的黄油中被发现的）、类糖物质（许多地球上的多糖的组成）及复杂的两亲物质（同时具有极性端和疏水端），它们是用氯仿从陨石粉末中抽提出来的，在水中能形成膜状囊泡[4]。

陨石和生物组分最有趣的相似之处是，一些手性氨基酸在陨石中呈L对映体过量（ee），这与在地球蛋白质中发现的类似，但是不像地球蛋白质过量那么多。

到目前为止，已经明确证实，0～15%的分子ee存在于α-氨基-α-烃基氨基酸这一亚类中[5]，是生物圈外唯一不对称的（细节见第7章）。在陨石氨基酸中发现的ee似乎说明，手性不对称性在前生物化学反应中起了部分作用。这些ee的起源还不知道，可能是在手性氨基酸形成时或形成之后，不对称性影响的结果。举例来说，圆偏振光（CPL）是辐射的一种形式，其电磁矢量绕转动轴旋转。这种旋转可以是顺时针，也可以是逆时针的；这样将产生具有手性的、彼此为镜像的右-或左-圆偏光，并将与手性分子的两种对映异构体产生不同的相互作用（见第7章）。如果这种辐射是“单手的”，能量足以破坏分子

（像紫外 CPL 一样），这两种对映异构体将以不同的速率被破坏，在暴露后[c]，一种对映异构体剩下的量将比另一种多。有人提出，这样的效应可能在陨石氨基酸形成时发生，即可能在到达母体小行星分子之前，在前体分子阶段就形成了（见后述）。

6.3　陨石有机物的复杂宇宙历史

总的来说，陨石有机组成的分析告诉我们，无生命的合成反应可以产生大量具有显著复杂性的有机物，包括与地球生物分子完全对应的组分。它们是如何形成的这个问题，具有重要的太空生物学意义，因为其答案将告诉我们，地球外无生命的合成体系可以产生生物分子，而这个信息将使我们能够阐明有机分子在宇宙中的分布，以及它们对地球及别处生命的起源可能具有的重要意义。

我们已经知道了陨石中复杂多样的有机物分子组分，但仍然未知它们的起源，目前正在对它们进行研究。陨石氨基酸可能是研究最为详细的，因为它们与地球上的生物多聚体（蛋白质）是对应物，并有 ee，它们被传递到早期地球的可能具有重要前生物意义。这些组分形成的机制可能也一样复杂。α-氨基酸有可能是由醛、酮、氨水及 HCN 在小行星母体暴露于液态水时合成的（框图 6.2）。

$$R_2-\overset{\displaystyle R_1}{\overset{|}{C}}=O + HCN \rightleftharpoons HO-\underset{\displaystyle R_2}{\underset{|}{\overset{\displaystyle R_1}{\overset{|}{C}}}}-C\equiv N \xrightarrow[NH_3]{H_2O} H_2N-\underset{\displaystyle R_2}{\underset{|}{\overset{\displaystyle R_1}{\overset{|}{C}}}}-COOH$$

框图 6.2

这种合成不能直接生成不对称的产物或形成非 α-氨基酸，我们还需要对可能途径进行假设。至于 β-氨基酸，有人提出，它们可能由氨水加入不饱和的腈中产生双键而形成，就像从丙烯腈中产生 β-丙氨酸（框图 6.3）一样。

$$CH_2=CH-C\equiv N \xrightarrow{NH_3} NH_2-CH_2-CH_2-C\equiv N \xrightarrow{H_2O} H_2N-CH_2-CH_2-COOH$$

框图 6.3

陨石有机物同位素组成的研究，使我们更加了解陨石组分的形成过程，因为分子中的同位素分布是它们合成历史的很好的指示物。

同位素是一种元素的不同质量的形式。所有元素的原子核都由质子和中子组成；质子，数量与电子相同，决定了这种元素是什么，但一种元素，可能含不同数量的中子，因而原子质量不同。所有生命所必需的元素（C、H、O、N），都有稳定的同位素，通常以在元素左上角标注分子质量表示。它们是^{13}C、^{2}H（它有自己的名字和符号，为氘和 D）、^{15}N、^{17}O 和^{18}O；也有同位素^{3}H（氚）和^{14}C，

但它们都有放射性衰变。

一种元素的同位素，具有不同质量，与其他元素或同位素成键时释放的能量不同。因此，在键断开或形成过程中（如 H 和 C 之间），这些元素的同位素在产物中的分布可能与在反应物中不同，这取决于反应的类型和条件。这种分布的改变来源于基于质量的同位素分馏，因此分子的同位素组成常常可以提示它们的合成过程[d]。同位素分馏是通过重、轻元素之间的比例（D/H、$^{13}C/^{12}C$ 等）进行衡量的；它与同位素之间的质量差异成正比，与反应温度成反比。较低温度下的 D/H 应当是最大的，因为 D/H 是稳定同位素质量比中最大的（与 C 的 13/12≈1.08，N 的 15/14≈1.07 相比，D/H 的 2/1=2）。

对这个预测的惊人证明，来自对星际介质（ISM）中浓密的云的分光测定说明，其中的分子富集了大量的 D。这些云由一大片气体和尘埃组成，如其名字所指出的，它们在星体间被重力松散地聚集在一起，可以延伸至几光年远。在百万年间，ISM 由已死星体的产物形成，并从这些云中不断形成新的星体。因为这些云中也进行着化学反应，所以 ISM 也是宇宙中化学演化的焦点，一种宇宙循环的中心，在这里，不仅每一代的恒星都从同一来源的不同元素池中产生，而且星际介质的元素、同位素和分子组成也在不断地演化（也见第 7 章）。

ISM 的温度能达到极低（5～10K），元素间（主要为氢，但也有氧、氮、碳和硫）的反应仅限于那些不需要活化能并且放热的反应，如离子反应和分子云中的气相分子反应。宇宙射线可以显著驱动这些反应，它能产生离子，并遵循之前提到过的分馏规则：同位素之间能量差别越大，温度越低，这种条件下产生的有机分子中，重的同位素越富集。实际上，在 ISM 中，观测到的 D/H 比例比地球上高很多（例如，星际间 DHO/H_2O 为 0.01，而地球上的平均值为 1.5×10^{-4}）。

对陨石有机物同位素组成的测量显示，陨石成分也在不同程度上富含 D、^{13}C 和 ^{15}N。一些单个组分的 D/H 比值尤其高，与光谱测定到的星际间分子的比值接近（表 6.3）[3]。用星际间分子的富集来解释这些同位素数据，可以得到关于它们的形成的一个通论，如下所述。

星际母体假说认为，冰冻的小行星是在小行星带上形成的，合并了大量的挥发性组分、水，以及富含氘的星际间有机物。这些可能是源于同位素衰变，碰撞的星体的变暖，后来的液相化学反应产生了大量不同有机物，以及陨石的固体与水交错的特征。跟着进行的是挥发性物质的散失和不挥发物质的保留。

这个概括，与陨石有机组成的许多分子和同位素的特征相符合。已知行星的历程对某些分子基团的影响可能比其他基团大，某些组分如支链氨基酸，在星际间的形成可能更为先进。因此，基于单个有机组分（表 6.3）的同位素组合，我们可以进行较详尽的推敲。

表 6.3 Murchison 陨石中一些可溶有机化合物中的平均氘含量

化合物	δδD* 值/‰
α-氨基酸，直线型	1800
α-氨基酸，分支型	3200
β、γ 等氨基酸，直线型	760
β、γ 等氨基酸，分支型	1800
羧酸，直线型	650
羧酸，分支型	1900
二羧酸	1000

* 太空化学家通过 D/H 的丰度测量星际间分子的 D 含量，而生物化学家和地质学家较喜欢测量样本及标准物中 D/H 的不同比例，并以这种差异占标准的千分率表示。这个过程与测量百分比类似，即 $[(R_{sample}-R_{standard})/R_{standard}]\times1000$。这种测量称为特定元素重同位素的 δ 值，使我们能够估计更小的同位素间的差值，而不需要用繁琐的比例。标准的 D/H 指海水中的值，定义为 0。有了这一定义，前面提到的星际间水的 δD 值为 65 667，而所有测量过的星际间分子的值为 5400～38 440。在南极冰中研究得较少的 CR2 陨石，其中一种氨基酸的 δD 为 7200。

6.4 太空生物学问题

我们仍不知道生命是如何形成的。但关于生命起源之前及之后的化学过程，我们已经了解了很多，而且知道这两个过程都是渐进的。从古生物学者的记录和对大分子的比较研究，我们知道生命开始得非常简单，经过了长时间的进化，并且到今天仍在进化着。我们也知道生物元素具有很长的宇宙化学进化过程，经历了这个过程，它们通常呈分子形式，有时还相当复杂。因此，提出以下这个问题就非常合理了：即这两个进化过程是否相关？若是，那么陨石将是含有生命起源之前宇宙化学进化的最后容器了，并且最适于回答一系列关于它“化学”起源的问题。

陨石化学反应告诉了我们什么？通过比较表 6.1～表 6.3 中对应的基本生化数据，很容易看到在陨石和生物圈中，有机物的分布似乎基本上不同。陨石组分具有很大的多样性（常常具有所有的异构体组分），为随机合成的产物；相反，生物分子为严格的组成选择的结果（例如，如前面提过的，仅仅 20 种氨基酸组成了所有地球上的蛋白质），且具有功能上的特异性。若没有对分子选择的引导或协助，很难说得出这样多样化的“一锅”有机物，在前生物化学进化上会具有什么先进性。

形成类似细胞的包裹体，可能帮助了有用分子的形成，并且是一些分子进化论的基础[6]。但是，在陨石氨基酸中发现的 ee，很有可能代表了所有来自地球外或实验室合成的非生命物质样本中，我们最希望能看得到的前生物印迹。这是因为，生物多聚体的手性相同性是产生生命的必要特性（见第 7 章）。因此非常有

理由提出，这种发展对生命的起源和（或）进化很重要。ee 的发现仍远远未具有决定性，它自身更引出了一些问题：陨石中 ee 的范围有多大？它们是如何形成的？输送到地球的不对称分子如何有助于分子进化？

这些都是目前太空生物学研究的基础，也是我们建议的专题阅读[7]。前两个问题在前面已谈到。至于第三个，在早期地球的分子进化中，这种陨石中的非外消旋组分，可能带来一些优势：

（1）现今，通过陨石、微陨石和彗星碎片，地球外有机物质向地球的输送是很大量的（每年 10^7kg 碳），而在早期地球被撞击的时代甚至可能会有更多；我们可以推测，有大量有机物质会到达地球上的适宜环境。

（2）陨石上含 ee 的氨基酸在水中不会变为外消旋（第 7 章），而难以预见前生物过程中手性相同性的发展，因为一些必需分子，如氨基酸和糖类，在水中是容易变外消旋的。因此，由于它能够保持其手性不对称性，这些陨石组分在分子进化的早期阶段应具有一定的优势。

（3）我们已知氨基酸具催化作用，在简单的前体分子合成的过程中，可将其不对称性传递给糖类[8]；这个效应对散播和扩大手性的前生物途径可能会带来启示。

6.5 概要

对含碳球粒状陨石的研究具有重要的太空生物学意义，因为这些陨石对引起地球生命产生、分布和复杂性的无生命有机化学的分析提供了唯一的地外物质样本。因此我们也能够对这种外源物质是否曾推动了早期地球前生物分子进化进行评价。

注释

（a）1969 年，美国宇航局实验室正在为月球样本的返回做准备，并迅速发现了陨石有机物的重要性，虽然最初对这块陨石的分析仅被认为是“尝试性的”（月球样本上反而发现并不含有任何有机物）。

（b）光谱分析，是通过分子/分子系统对辐射的吸收或发出测量它们能量转移的技术。

（c）通过 UV CPL 进行的差异性光解，用各向异性因子 g 表示来自于两种对映异构体的不同 CPL 吸收比例（由它们的消亡系数表示）及平均总吸收。即这种吸收区别越大，分子的 g 越大，两种对映异构体光解的程度和产生 ee 的量相差也越大。g 是一个精确的物理参数，决定了手性分子在反应时间内，两种对映异构体完全分解之前，产生 ee 的程度。对氨基酸而言，g 相当低，为 0.02，限制 UV CPL 可能产生的 ee 最多约为 10%。

（d）举例来说，在寻找古老生命形式的化石记录时，是通过碳的低$^{13}C/^{12}C$比例来分析识别生命物质的，因为就我们所知，生命是相当懒的，在其代谢反应或产生生物高聚物时，更喜欢用较轻的^{12}C。

参考文献

以下为在本文中引用过的参考文献，它们提供了与标题有关的领域的额外阅读，同时专题文献也提供了著作中的分析性文章。

扩展阅读

1. Taylor SR. Destiny or chance, our solar system and its destiny in the cosmos. Cambridge University Press, Cambridge UK, 1998.
2. Meteorites and the Early Solar System II, DS. Eds. Lauretta and HY McSween Jr., University of Arizona Press, Tucson, 2006.

专题文献

3. Pizzarello S, Cooper GW, Flynn GJ. The nature and distribution of the organic material in carbonaceous chondrites and interplanetary dust particles. Ibidem 625-651.
4. Deamer D. Boundary structures are formed by organic components of the Murchison carbonaceous chondrite. Nature 1985; 317:792-794.
5. Cronin JR, Pizzarello S. Enantiomeric excesses in meteoritic amino acids. Science 1997; 275:951-935.
6. Bachmann PA, Luisi PL, Lang J. Autocatalytic self-replication micelles as model for prebiotic structures. Nature 1992; 357:57-59.
7. Pizzarello S. The chemistry of life's origin: A carbonaceous meteorite perspective. Acc Chem Res 2006; 39:231-237; and references therein.
8. Pizzarello S, Weber AL. Prebiotic amino acids as asymmetric catalysts. Science 1997; 303:1151.

7 手性、同手型与生物分子相互作用规则

Sandra Pizzarello

7.1 定义与历史背景

以前，手性被定义为物体不能经平移或旋转与其镜像重合的性质。这个单词来源于希腊语 χειρ，意思是手，实际上，类似的词——“手型”可以将这种性质解释得更为贴切。类似于手，手性物体有两种形式，互为镜像，组成一致，但是当被叠印到一起时却不能匹配，因为它的组成空间分布不同，试着将右手套入左手的手套或将右脚放入左靴，就能给出令人信服的证明。手性以多种二维或三维形式在自然界中广泛分布（图 7.1），且可以与运动相联系，例如，当一个物体进行螺旋运动时，像螺钉一样，它既可以从左到右，也可以从右到左。

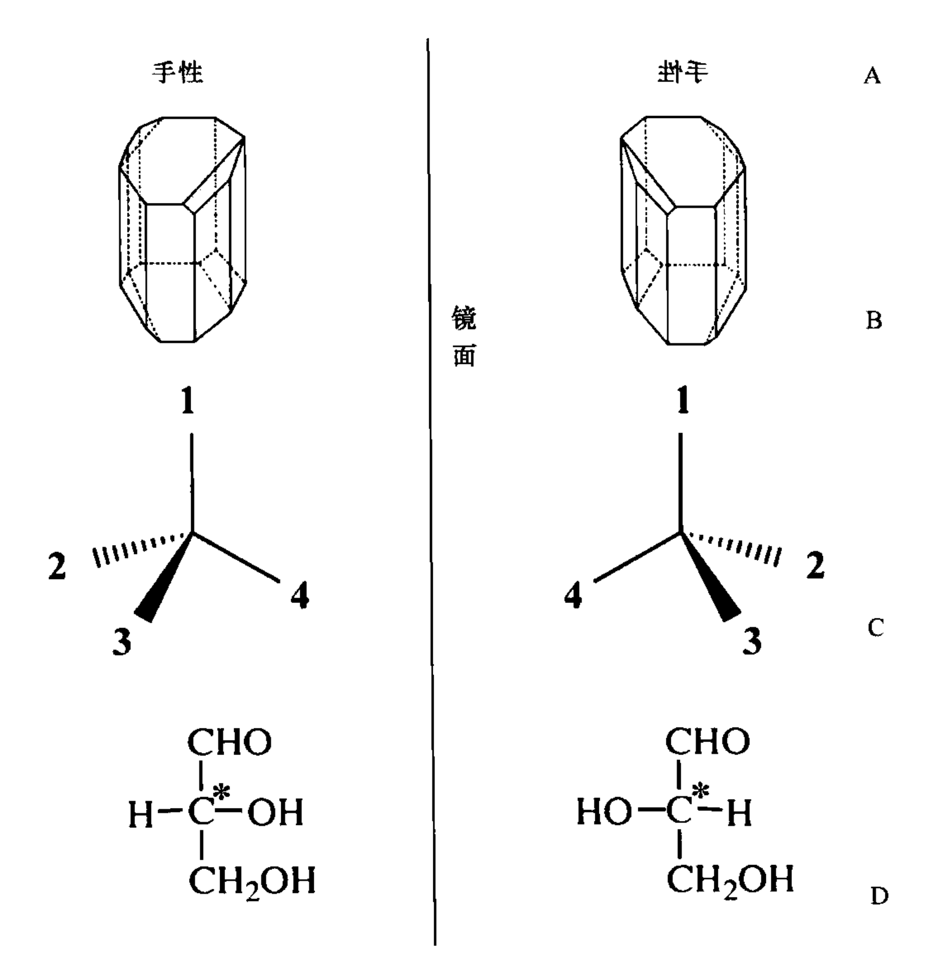


图 7.1　手性。

从镜子里看手迹是手性的一个很好的例子，也是如何发现它的很好提示（图7.1A）。轴对称或中心对称的字母（H、O、I等）在手迹或其镜像中是相同的（即非手性），而不对称的字母（C、L、R等）在镜子中看起来不同，为手性的。如图中显示的三维形式，酒石酸盐晶体（图7.1B）是手性的，因为它的不对称的表面分布不同。它具有重要历史意义，是由路易斯·巴斯德（1822～1895）发现的第一个具有手性结构的有机分子[1]。巴斯德很有天分，在许多领域都是先驱型的实验科学家，在他事业的早期，他分析了通常沉积在酒桶底部的晶体，它被称为酒石，已知是由酒石酸盐构成的。他发现在实验室合成时，这些盐可以结晶成如图7.1所示的两种镜像形式，但在酿酒过程中只能自然形成其中一种形式。这在最初似乎是矛盾的，因为这两种结晶具有相同的化学方程式，但巴斯德具有洞察力，在偏振光及其与物质的相互作用基础上设计他的实验，偏振光及相关现象是在17世纪和18世纪早期被发现的。

如果我们认为光是一种电磁场的波动，其振动平面互相垂直，且与传播方向垂直（图7.2A，分别为H和E），当光没有被过滤时，这种振动将在无数个平面内发生（图7.2B，沿这条线上的一点看）。但是，当光线穿过矿物方解石透明的晶体时，人们发现它将偏振到一个平面上（图7.2 C）。当通过特定的"旋光性"物质时，还发现这种偏振光平面能够旋转。通过对酒石酸晶体进行实验，巴斯德论证出，只有天然的、单手性的酒石酸晶体溶液能够旋转偏振光，而其两种形式的混合物却不行。在做这个实验的同时，他得出两个意义深远的发现：不仅只有晶体，有机物分子都可能具有两种空间结构，而地球上的生物分子只有单手型——这就是说，它们具有相同手性，或按后来的定义说，具有同手型。

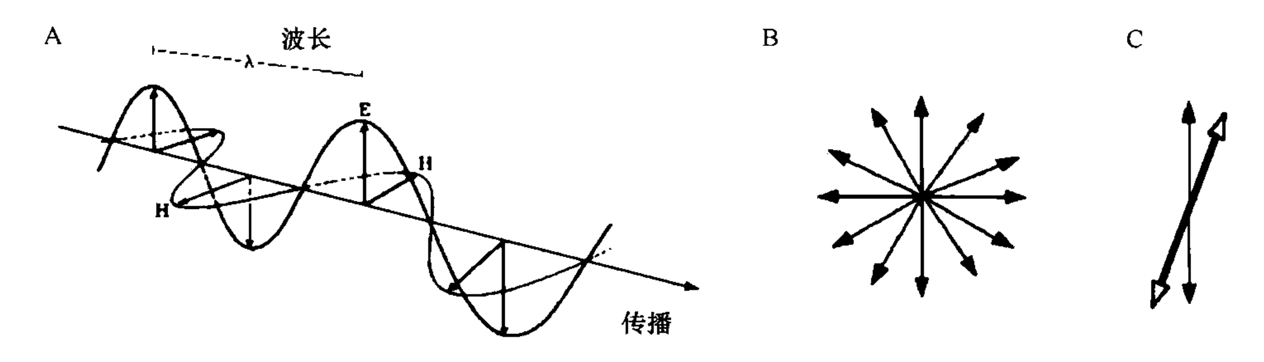


图7.2　光的电磁矢量及其运动和方向。

7.2　有机物的手性不对称与生物学的同手型

正如我们现在所知，实际上，手性最普遍且最有趣的例子由含碳（C）化合物显示出来。这种元素有4个结合价，其空间分布是这样：分子中的C可以被

看做是一个常见的四面体，C在中间，4个结合基团作为顶点（图7.1 C、D）。当所有这些基团都不同时，C成为非对称的（C^*），含这种碳原子的分子为手性的，即它们可以以两种“手的”形式存在，称为对映异构体。手性分子的对映异构体具有对应的化学性质，与非手性分子和结构作用时是等价的。但是，由于其手性碳取代基不同的空间分布（构型）和其电子的不同极化性等限制，它们不具有相同的物理化学性质。例如，手性分子的两种对映异构体的溶液将以相反方向旋转偏振光，只有当它们的量相等时，才可以抵消彼此的影响，溶液不旋光（这种对映异构体的等量溶液被称为外消旋的，由巴斯德对合成酒石酸的分析得到）。图7.1D给出了甘油醛的Fisher投影式，甘油醛是最简单的手性糖类，其对映异构体习惯上定义为D型和L型，分别来自对偏振光的右旋和左旋。在定义其他糖类和氨基酸的对映异构体时，也把甘油醛作为参考物。

对映异构体间有一个不同点，由上述性质衍生出来，且对于生物化学最为重要，即它们与其他手性分子的反应是不同的。即使是相同化合物，对映异构体遇到对映异构体的相同或相反构型时，其行为不同。举例来说，L型氨基酸与其L型或D型对映异构体的相互作用/反应不同，其DD和DL复合物的能量不同。

再次用手进行类比可能有所帮助。若不考虑它们通常在长度上的不同，左手和右手在拿起一个对称物体（如球）的时候作用是相同的。但是，当面对一件手性物体时，如另一只手或者一个旋转的阀门，左手和右手就会对这些物体的两种手性形式感觉出不同的难易程度。

因此有必要说，手性是分子及物体相互作用的性质，正因为如此，它成为生命分子的必要条件。由于现存的生物都基于碳化学[a]，以及碳能与自身或其他原子成键的独特能力，在组成生物圈的无数有机化合物中，许多化合物都具有至少一个非对称碳。而且，如巴斯德发现的，现存生命对手性化合物有精确的应用，用单手型的单体，如L型氨基酸和D型糖构建高聚物，如蛋白质、RNA、DNA、多糖。许多代谢活动也依赖于精确的手性相互作用，结果是手性构型决定了我们体内及与环境之间的许多化学通信，后者包括药物治疗的效果。关于药物治疗有无数的例子：有悲剧的，在20世纪60年代早期，当一个怀孕的女人只吃下了镇定药物的一种对映异构体时，导致了胎儿的不正常发育；也有有趣的，比如化合物萜二烯的R型闻起来像柠檬-橘子，而S构型的对映异构体则像菠萝[只有糖和氨基酸使用D和L标记，而所有其他有机化合物都用R型和S型表示，来自拉丁文rectus（右）和sinister（左）[b]]。

生物分子的这种手性相同性，称为同手型（homochirality），就我们所知对生命是必需的。如果我们认定生命依赖于其高聚物的功能（见第6章），那么就很容易理解这种必需。这是由于生物高聚物的活性依赖于其空间结构，而它超结构的功能特异性由其单体分子的同手型决定。现存生物中，蛋白质和DNA的螺

旋，蛋白质可能形成的折叠，以及底物和酶的锁定结合，都要归功于单手性的单体。若生物高聚物以不同的手性形式组成，生命还能行使其功能吗？例如，若不是同手型，D 型和 L 型氨基酸能共同构成蛋白酶吗[2]？只要蛋白质序列仍以精确、可复制的方式被决定，正如现在的酶中的 L 型氨基酸一样，那么我们应该可以假设，D 型和 L 型单体也能够产生一些精确和合适的空间结构，具有能够行使功能的，与靶分子匹配的活性位点。比起现在的同手型结构，构建这样的结构可能更不经济，但也没有理由认为它们会破坏生命系统。但是，由 DNA 给出的基本例子来看（框图 7.1），事情并不总是这样。这种生物高聚物由双链分子拧成螺旋构成，每条链含 D 型糖磷酸重复骨架，其上连接着精确互补的 4 种碱基——腺嘌呤、胸腺嘧啶、胞嘧啶和鸟嘌呤，这些碱基在双链间彼此以氢键连接，维持螺旋的完整性。改变沿链任何一个糖的手性都将破坏这种特定次序和互补，没有形成双链结构的可能。由于生命过程的基础与 DNA 相关，如转录和复制，有赖于这种特定结构的维持，它的任何改变都将同时妨碍它的功能。因此，以我们对生命的了解，我们不得不得出结论，对地球生物分子而言，同手型是个必要特点。

框图 7.1　DNA 和 RNA 糖化合物的同手型使得碱基能够配对，并在链之间互相交换氢键，维持螺旋的完整性。

7.3　对地球上同手型起源的假设

这种同手型是怎样起源的？同手型对生命的起源和（或）发展是不可或缺的

吗？它开始时是广泛存在还是有限存在的？这些太空生物学问题都是自从巴斯德以来大家就开始争论的，并且与生命起源相关的问题一样，仍然是个谜。关于生命和非生命的理论都已被提出[3]。关于生命的理论认为，生命从非手性的和（或）外消旋分子形成，经历了时间过程，手性选择发展并进化了；而非生命的理论认为，在地球生命产生之前，转向同手型的趋势在非生命化学进化中是固有的（在陨石氨基酸中发现了 L 型对映体过量，这支持了后者，尽管没有证明——见第 6 章）[2,3]。

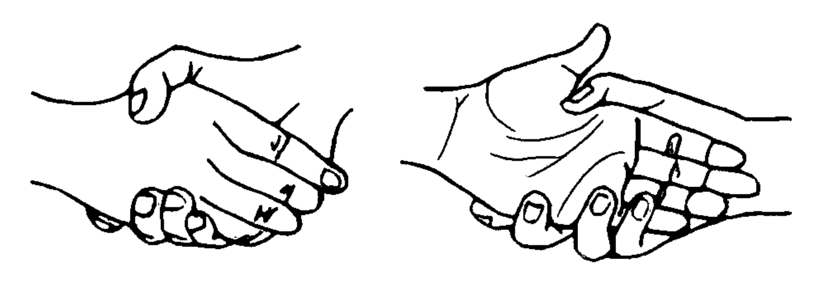


图 7.3 “正确”和“错误”的握手。

关于生命的理论天生就难以证明，这是因为围绕生命起源及养育它的早期地球环境，都具有不确定性。要提出一个向同手型进化的途径，其基本原则是，氨基酸和糖在水中都很容易外消旋[c]，而初始时少量的对映体过量极易丧失，除非生命本身发展出了产生手性相同性的机制[c]。在这里应该提到，带有对映体过量的陨石氨基酸不会外消旋，是因为它们缺乏与羧基相邻的H，而它在水中的移除和再插入将导致外消旋的产生。已经用非手性核酸类似物建立了一个实验室模型，可以线性携带一列碱基并形成螺旋结构[6]。

非生物理论面临着另一个问题：除了已提到过的陨石氨基酸中的对映体过量，在地球生物圈之外还没有测出过分子不对称性。因而，前生物起源的同手型假说就得要么依赖于陨石和彗星对早期地球的大量外源物质的传递（见第 6 章），要么就假设，既有对称性的破坏，又有这种内在不平衡性的扩大。后面这组假设需要得到决定其偶然性或必然性的内在原因，其中一些已在下面的段落中列出。

最吸引人且显然最普适的想法是：手性分子的对映异构体之间有一个内在的能量差，这在理论和分析研究中已经提出过了。如前所述，手性分子有相同的化学性质，即由电磁力驱动的分子的运动和相互作用，不因空间映像的不同而不同，其镜像发生反应的可能性是相等的。但是，有人发现[4]，这对分子的原子内水平上所有已知作用力都是不适用的，如重力、前面提到的电磁力、强核力（已知最强的力保持原子核为一整体）和弱核力。

在弱核力中较弱的力里，发现了手性趋势：其最广为人知的效应就是在放射性衰变中产生的 β 射线，这些射线里含左旋的电子比右旋的要多。

这个发现被解释成[d] 分子的对映异构体不能具有完全相等的能量，这产生出

两类假设。①通过化学进化中无数的非生物和前生物反应，对映异构体间的能量差可能被扩大，最终导致了同手型。这个假设的缺点在于其理论上预计的能量差可能极其微小，尚且没有对任何一种特定分子得出证明，无论在理论上还是分析上都没有达到令人满意的共识。②β衰变中产生的电子可能影响前生物分子合成中的左手趋势（见后）。

虽然不普遍，但还有一些高频的例子，其中对映异构体能够自发分离，然后经历不连续的变异过程，使一种异构体相对另一种有所增加。结晶过程提供了很好的例子。基于影响它们形成过饱和溶液的物理化学条件，手性化合物可能包含两种对映异构体，也可能只含其中一种的晶体（如巴斯德在酒石酸中的发现）。也有一种晶体（如石英）其手性不是因为组成分子（SiO_2）的不对称，而是由于其分布朝向的不对称。

有人提出，在局部环境中，一种晶体类型偶然产生的主导地位，可能导致手性趋势的发展，如对其他手性分子的吸收（如氨基酸）。我们也知道，在结晶过程的初期阶段，形成的早期手性晶体（再次声明，这是偶然的）可能在溶液中为之后的结晶过程“成核”，导致产生一组基本上同手型的晶体[7]。请记住，偶然事件可能对每种对映异构体都是同等有利的，并将导致局部环境中的每种手型在整体上随机分布（而最终其中一种形式将“击败”另一种形式，造成整体上的同手型）。

其他可用来增加消旋手性化合物异构体的效应，是基于它们与手性结构、分子和力的相互作用。例如，前面提到过的黏土和晶体（如石英）的非对称吸收；多聚化过程中的动力分割，利用同手型和异手型间的能量差，促进同手型的形成（见7.2节）；以及与圆偏振光和β射线的作用。虽然大多数基于这些效应的假设与偶然性有间接联系（等量的D型和L型石英晶体，或等量的D型和L型蛋白质等），手性力的效应可能与广泛的非生物环境相关性更大。我们知道β射线具有普遍的适用性，而圆偏振光亦在大片的星际介质中被观测到（见第6章）。

介于生物和非生物理论之间，还有一种“泛种论”，字面意义是万有的来源。它认为生命是由太阳系外的天体带到地球的（由太阳风帮助，最初由Arrhenius于1908年提出）[5]。如果这些“种”或孢子具有同手型组成，它们将解决地球上同手型起源之谜。但是，更具普遍性的谜题仍然存在，因为“泛种论”将手性及其他有关生命起源的问题推到了另一个时空。

7.4　概要

手性或手型，为缺乏中心对称性的物体及分子，它们具有组成相同，仅在空间分布上有所区别的性质。大量的含碳分子都是手性的，其内在的非对称性，为现存生命同手型大分子的构建、结构和高聚物功能提供了相互作用方面的有利条

件。这种手性相同性对地球生命是必需的，而地球生命的开始可能也有助于同手型的发端。进行同手型来源的探寻对太空生物学具有重要意义。

注释

（a）有机物一般是指碳化学，但实际上，该名词是来源于 18 世纪的概念，即只有活的生命体系才可以产生这些体系所含的化合物。

（b）有机化合物的绝对构型具有“程序法则”：①C^* 上的 4 个取代基的级别（图 7.1D）由其原子序数确定（序数越低，级别越低）；② 对手性分子进行象征性的空间定位，只要从级别最低的碳键向下看［将看到剩下的三个取代基以不对称 C 原子（C^*）为中心呈轮辐状辐射］；③ 现在若将取代基由最高级向最低级连线，其路径逆时针方向的为 S 型（左，sinister），顺时针方向为 R 型（右，rectus）。

（c）由于电子向邻近的氧原子偏离，在水中倾向失去一个 H^+，临近羧基的碳原子具有微弱酸性，形成的离子可以再获得一个 H，这个过程可能导致构型改变。经过一定时间和合适的条件，这个过程会导致外消旋混合物的产生。

$$^{+}H_3N\text{–}C(CO_2^-)(H)\text{–}R \;\underset{+H^+}{\overset{-H^+}{\rightleftharpoons}}\; ^{+}H_3N\text{–}\ddot{C}^-(CO_2^-)\text{–}R \;\underset{-H^+}{\overset{+H^+}{\rightleftharpoons}}\; H\text{–}C(CO_2^-)(NH_3^+)\text{–}R$$

（d）对对映异构体间能量差的解释（称破缺能量差或 PVE），来源于一个统一了电磁和弱力的物理理论。由该理论可知，这个力能够同时与核和电子相互作用，产生“左”和“右”的区别[4]。

参考文献

扩展阅读

以下为在本文中引用过的参考文献，它们提供了与标题有关的领域的额外阅读，同时专题文献也提供了著作中的分析性文章。

1. Geison GL. The private science of Louis Pasteur. Princeton University Press, Princeton, New Jersey, 1995.
2. Cronin JR, Reisse J. Chirality and the origin of homochirality. In: Gargaud, M. ed. Lectures in Astrobiology. Vol. 1; Berlin: Springer-Verlag, 2005:473-514.
3. Bonner WA. The origin and amplification of biomolecular chirality. Origin Life Evol Biosphere 1991; 21:59-111.
4. Hegstrom RA, Kondepudi DK. The handedness of the universe. Scientific American 1990:108-115.
5. The web-site: www.daviddarling.info/encyclopedia/P/panspermia.html from "The Encyclopedia of Astrobiology, Astronomy and Spaceflight" offers a descriptive entry on the subject of Panspermia.

专题文献

6. Nielsen PE. Peptide-Nucleic Acid (PNA)-A model structure for the primordial genetic code. Origins Life Evol Biosphere 1993; 23:323-327.
7. Kondepudi DL, Kaufman RJ, Singh N. Chiral symmetry breaking in sodium chlorate crystallization. Science 1990; 250:975-976.

8 早期地球

James F. Kasting

8.1 引言

要理解生命起源，就有必要了解早期地球的环境，即使是“泛种论”的观点——生命是在其他地方起源之后转移到地球，这个结论也仍然成立。就算按照这种有些极端的观点，要想预测生命到达这里后是如何适应的，也需要去了解地球早期的环境是什么样子的。较传统的模型认为生命在地表某处起源，那么对地表的了解就更为重要了。在这里，我们回顾一下目前关于地球及其大气和海洋起源的认识，然后讨论关于早期大气的组成和环境，有哪些是已知的，哪些是未知的，最后，我们简要涉及早期生命对大气的影响。最后这个话题与前生物进化没有直接关联，但是它与全局相关，因为探索生命起源的最好方法之一就是追溯到最早的有机物，而要做到这点，我们需要知道它们与其所在的世界是如何相互作用的。

8.2 地球与其大气、海洋的形成

8.2.1 历史回顾

对研究生命起源的学者来说，早期大气和海洋的性质一直是他们感兴趣的话题，关于这个问题，很早就有的一些看法仍然受到重视。我们可以从俄罗斯生物学家 A. I. Oparin 的书开始，这本书最早于 1924 年在俄罗斯出版，后来在 1938 年有了英国版[6]。类似的观点来自 1929 年 J. B. S Haldane 在英国出版的书，因此后来关于生命起源和早期地球性质的模型常被称为“Oparin-Haldane”假说。在这个模型中，太阳和另一颗恒星的近距碰撞喷射出的过热气体凝缩而形成了早期地球——这个机制现在认为基本是不可能的。这个模型预测，地球最早的大气是高度还原性的，也就是说，它富含氢气（H_2）、甲烷（CH_4）、氨气（NH_3）。在这些气体中，任何一个都能与氧气反应。在 Oparin 模型中，早期大气温度很高，大气中的热能使这些还原性的气体相互反应，形成一层厚厚的浆状物，或有机化合物的“汤”，认为生命就是在其中出现的。

虽然 Oparin 关于地球形成的理论后来看起来不能让人信服，但是他关于初期高度还原性大气的模型几年后还是得到了支持，即来自芝加哥大学的地球化学家 Harold Urey 及他即将毕业的学生 Stanley Miller（在写这一章期间，米勒不

幸去世了。我们这些认识他的人将会非常想念他)。Urey 是一位知名的地球化学家,他的研究兴趣包括对木星和土星的光谱分析。Urey 从这些观察证明了这些大行星的大气中富含甲烷和氨气。他也提到,这些行星非常大且重,它们实际上保留了其形成过程中获得的所有原始氢。相反,较小的行星,如地球、金星和火星,则将氢散失到空间中,这使得它们的氧化性随时间而增强,Urey 推论说,在地球还未失去氢之前,其大气应类似木星和土星。因此,当他的毕业生米勒提出研究放电作用(模拟闪电)对这样的大气混合物的影响并把它作为博士论文时,Urey 让他尽管去做。剩下的故事就广为人知了:米勒的实验得到了大量令人信服的前生物有机化合物,包括氨基酸和核酸[7],并且"Miller-Urey"关于生命起源的模型在科学文献和流行科学意识中都占据了稳固地位。

但是,几乎与此同时,另一位地球化学家 William Rubey 建立了地球大气起源的不同模型[8]。Rubey 研究了现代火山口散发的气体的组成,他发现它们主要由 CO_2 和 H_2O 组成,而不是 CH_4 和 NH_3。如果如所想的一样,地球大气由火山口放气形成,如果火山气体组成保持不变,那么最早的大气应被 CO_2 和 N_2 主导,而不是 CH_4 和 NH_3 (N_2 在现代火山口放气中很难被测到,这是因为很难将火山口的 N_2 与空气进行区分,但 Rubey 的理论预测,碳和氮都应以较氧化的状态被放出)。

Rubey 的理论后来被其他工作者详细阐述,尤其是 Heinrich Holland 和 James Walker。Holland 建立了一个二阶模型,认为火山气体最初还原性较强,后来在地核形成后,氧化性较强。我们现在确信,地核是在星体增长过程中形成的(见下一节),因此 Holland 的高度还原性的"第一阶段"可能并不存在(或者即使它存在也是因为不同的原因)。Walker 在他的论文和书中[1] 建立了现在称为弱还原性的早期大气模型。在他看来,早期大气主要是 CO_2 和 N_2,和 Rubey 提出的一样,但至少在地面上仍缺乏自由的 O_2。分子氢 H_2 以 100~1000 ppmv (百万分单位体积)的浓度存在。将火山口放出的微量氢气和逃逸到空间中的氢进行抵消,Walker 估计了 H_2 的丰度。这种平衡大气氢收支或氧化还原收支的想法,对了解早期大气组成是非常关键的,我们在下面将回到这一点。

如果 Walker 是对的,早期大气只有弱还原性,那么我们关于生命如何形成的观点就需要改变了。Miller-Urey 类型的前生物化合物合成在这样的大气中效率将非常低[9]。在弱还原性大气中,可能通过光化学合成小分子,如甲醛 (H_2CO) 和氢氰酸 (HCN)[10,11],但是形成复杂有机化合物,如氨基酸和核酸的速率将非常低。意识到了 Miller-Urey 模型关于生命起源的这个问题,许多研究者提出了不同的模型。在这些模型中,非常突出的是"喷汽孔"模型,认为生命起源于洋中脊热液循环系统中合成的还原性有机化合物,以及"空间起源"模型,认为复杂有机物来自于陨石和星际间尘埃粒子 (IDP) 的组分。

8.2.2 通过小行星体的撞击放气形成大气

我们对早期大气组成的了解有多正确？我们真的知道早期大气就是弱还原性的吗？而且，如果真是这样，这又意味着什么？

对这些问题的确切回答仍非常不明确，这是因为地球的形成是一个复杂的过程。也许最重要的是，大气最初可能不是来源于火山的放气。在 Rubey 的时代，人们认为地球是小行星体的缓慢增长（accretion）形成的（增长是用来命名环绕年轻恒星的小固体天体彼此碰撞，黏到一起形成较大天体的过程）。我们现在相当肯定，不是这样的。实际情况是，地球形成得相当快，在几千万年间就形成了，而其增长后期，牵涉了与月亮大小类似，或者甚至更大的小行星体的撞击（月球的质量约为地球的 1/80）。实际上，月球自身的形成，就可能是因为火星大小的天体，质量至少为地球 1/10 的天体撞击形成的，也可能是这个质量的两倍。

巨大的小行星体撞击一颗形成中的行星表面，会释放相当可观的能量，其中一些能量将用于熔化地球表面。实际上，有确凿的地球化学证据，能够证明地球早期历史中存在过一个或多个岩浆的海洋。这样广泛的熔化可能使金属铁从熔化的硅酸盐中分离出来，从而渗透下去，形成地核（我们现在之所以认为核的形成与行星的形成是同时的，这是原因之一。由核形成的多种地化指示物，如钨的同位素[12]，可得到类似结果）。撞击天体散失的能量可能更大，因此较大的小行星可能在撞击时被蒸发掉了。它们所含的任何挥发性物质将因此直接注入大气（挥发性物质为一些化合物，如 H_2O、CH_4 和 NH_3，具有低沸点，因而很容易蒸发），这个过程被称为撞击放气。撞击放气在地球大气和海洋形成的过程中，几乎一定起了重要作用。但是，海洋并没有马上形成。由于地球表面极热，水最初应处于蒸汽形式，形成了浓密的蒸汽大气，其表面压力比现在高 10 倍，甚至 100 倍[13]。

但大气形成的过程是更加复杂的。许多挥发性化合物，包括水，在硅酸盐熔浆中是可溶的，而且若原始地球很小，到来的撞击物是不会被完全蒸发的。这样，一些水及其他挥发性物质将加入形成中的地球，后来被火山放气释放出来。在伸展的洋中脊中，观察到有原始的（非放射性成因的）^{3}He 被放出，证实了这个猜测，因此，Rubey 至少是部分正确的。此外，大撞击还可以通过直接把大气吹到空间中将它移除，这个过程称为冲击侵蚀。要确定撞击物是否促进了挥发性物质的积累或散失是一个难题。举例来说，形成月球的撞击可能吹掉了地球所有预先存在的大气和海洋，需要整个形成过程重新开始。令人惊异的是，如果撞击发生的时候海洋处于凝聚态，吹掉的过程将最为有效，这是因为撞击产生的冲击波将有效转化为能量，蒸发掉水，使过热蒸汽将其上的大气气体吹走[14]。

撞击产生的大气的组成是什么呢？在高温下，撞击蒸发产生的气体将由小分子组成，如 H_2、H_2O、CO、NO 和 N_2。但随着撞击云的降温，较大的分子将会更具热动力学优势。撞击云自身将成为高度还原性的环境，可能足以使碳转化为 CH_4[15]。这对于那些被埋得很浅而没有蒸发的较小撞击物也是成立的[16]，可能产生的 CH_4 的量最初大得惊人。现在地球表面碳的总量，包括含碳岩石，约和金星一样，足以产生分压为 60～80bars 的 CO_2（金星的 CO_2 分压为 90bars）。如果这些碳最初都多数以 CH_4 存在，CH_4 分压理论上应能达到 20～30bars（与 CO_2 比较，CH_4 分子质量较低，需约除以 3 来进行补偿）。这种估计差不多是偏高的，因为部分来自撞击物的碳最初会形成 CO 和 CO_2，而且 CH_4 很快就会经光化学反应开始消耗，其中一些将被氧化为 CO 和 CO_2，而另一些可能聚合形成更高级的碳氢化合物。即使这样，似乎很明显的是，随着星体增长和其后的大撞击，仍有大量的 CH_4 会保留下来。

这将给我们带来一个难题：在地球的早期历史中大撞击的频率是多少？如专题 8.1 中所述，我们不知道这个问题的答案，因为这取决于还没有解决太阳系最初的几亿年间陆地行星所经历的碰撞率。因此，关于那时大气的情况，很难得到确切的结论。在下一节中，我们将概述什么是我们所知道和不知道的。

专题 8.1　大撞击是持续的，还是一次性爆发？

自从 NASA“阿波罗”人类登月计划（1969～1973 年）以来，科学家们就知道，月亮上许多岩石的年龄为 3.8～3.9 Ga，这最初被解释为撞击物的一次爆发，它们都在那段时间附近抵达，被称为晚期大冲撞。休斯敦月球及行星研究所的 Graham Ryder 是这个理论的带头支持者，他坚持着这个观点直到去世[17]。但是，由于已知地球的膨胀是在 4.4 Ga 左右或更早的时间进行的，后来的理论家们很难理解这样的撞击发生于地球进化相对较晚的时期（U-Pb 对陨石进行的年代测定，对早期太阳星云的年龄一律给出了 4.55 Ga 的结果，我们认为地球通过增长作用的形成是在那之后一亿年内完成的）。因此，修正主义者认为，在 4.5 Ga 和 3.8 Ga 的整段时间内，存在一个晚重撞击期，期间在主要的增长时期之后，撞击流以指数方式逐渐减小。如果是这样，那么这种不断的、偶发的外来物体流可能在整个早期历史中会影响到大气的组成。

但是，在最近两年，一个新的太阳系形成的动力学模型为原先晚重撞击期的“爆发”理论提供了支持[18,19]。这个模型有时被称为 Nice 模型，因为该理论的一些工作者的实验室位于法国南部 Nice 市附近（这个城市的名字发音为“neese”，而不是“nice”）。这些作者利用复杂的天体机制编码来模拟行星增长过程的后期阶段。在模拟开始时，他们利用一群基本上平均分配的月亮大小的行星体，然后计算它们彼此之间的重力作用和它们形成行星时的碰撞。4 颗大行星——木星、土星、天王星和海王星，被认为在模拟开始时完全形成，但是它们所处的位置与现在的轨道半径不同。在模型中，木星形成的位置距太阳比现在略远，而土星则略近。天王星和海王星被认为是紧挨着土星轨道的，海王星距太阳比天王星更近（这与今天的情况相反）。所有这些假设都是可能的，即使它们不可能是形成现在太阳系的唯一起始点。

模拟得到的结果很有趣，并且可能与实际上发生的非常接近。由于与较小的行星体之间的重力相互作用，在这个过程中木星向内迁移，而土星向外迁移。在 3.9 Ga 左右的某一时段，土星与木星经历了 2∶1 共振，即土星的轨道周期恰好为木星的两倍（这只比现在略大）。这种共振促使这两颗行星进入高度偏心（椭圆）轨道，而它又反过来影响了两颗质量较小的大行星——天王星和海王星。这两颗行星都被推离太阳，海王星从天王星轨道内部运动到外部。尽管这个复杂的过程在某种程度上可能看起来有点特别，但它是对初始动力系统应用牛顿运动定律的直接结果，并与我们对有关大行星迁移的知识一致，这些知识是通过对太阳系外行星的研究得到的。模拟的最终结果是，天王星和海王星突然被抛到外太阳系，而那时外太阳系还充满了冰冻的小行星体，它们还没有足够的时间去增长为更大的天体。后来这些小行星体多数都分散离开了原先的轨道，其中一些穿过内太阳系，并有一些撞击了月球和类地行星。如果这个过程或类似过程是正确的，那么最初对月球岩石的解释可能也就是正确的，即它们意味着一次大撞击的爆发。

幸运的是，这场争论可能会在不远的将来得到解决。NASA 已经计划在下一个 10 年内派宇航员重返月球。通过从更广泛更多样的地点收集月球岩石，应该有可能得到这个问题的答案，确定大撞击到底是持续不断的，还是相对突然的灾难性事件。

8.3 前生物大气与环境

前面的讨论为描述前生物大气的组成做好了铺垫。很明显，由于前面提到的那些不确定性，尤其是地球的碰撞历史，不可能提供非常确切的答案。但是，我们仍可以尝试给出一些结论，列出如下。

(1) 一个凝缩的、高度还原性的大气主要由 CH_4 和 NH_3 组成，而其他还原性气体的可能性不大。尤其是 NH_3 可能含量不高，因为它可以迅速光解，转化成 N_2 和 H_2[20]。氢逃逸到空间中，留下了稳定的、三键的氮气作为主要的含氮气体。碳氢化合物形成的薄雾对紫外线的防护也不能够避免这种情况的发生[21]。的确，这个结果取决于粒子大小的情况，而随着更成熟的薄雾模型的建立，这种情况是可以改变的（更小的粒子可以产生更好的紫外线防护作用）。但是目前，这个模型显示，Miller-Urey 类型的合成在产生前生物有机化合物上，不是一种有效方法。

(2) 除了第一点，大气也有可能富含 CH_4、CO 和 CO_2。撞击会为 CH_4 提供大量的非生物来源，尽管这一点还需参考讨论地球撞击历史的专题 8.1。CH_4 的另一个可能非生物来源是富含 CO_2 的海水与超镁铁质（富含铁和镁）岩石的相互作用，这个过程称为蛇纹石化作用，在洋中脊热液循环系统中发生。目前，我们认为 CH_4 的这个来源是有限的，在缺氧的地球上，产生的大气 CH_4 浓度仅能达到百万分之几[22]。但是在过去，如果由于更高的地热能流动及板块构造的改变，使更多的超镁铁质岩石暴露，CH_4 产量可能更大。这很难定量，因为我们不能真正知道在遥远的过去板块构造是怎样的。含 CH_4 和 H_2O 的大气，必然也

含有一定量的 CO 和 CO_2，因为甲烷会被水蒸气光解生成的副产物氧化。由于氢向空间的逃逸和 CH_4 不能有效地通过光化学反应再形成，这个过程是半不可逆的。CH_4 的再生需要表面相互作用，如前面提到的蛇纹石化作用。如果大气中的 CH_4/CO_2 比例大于 1，一些 CH_4 也可通过光化反应转化为更高级的碳氢化合物。在生物出现之前的地球，不知道是否达到这个比例。在生物出现后的地球，似乎已达到了这个极限，形成了碳氢化合物的霾[21]。

(3) 正如之前提到的，在生物出现之前的地球，大气中的 CH_4 含量很可能是有限的（10～100ppmv），并也能通过光化学反应产生 HCN 和 H_2CO。这些化合物对氨基酸、核酸和糖类合成的出现是非常合适的，虽然我们尚且不了解这些复杂化合物形成的细节。由于其热动力自由能非常高，CO 也有理由大量（100ppmv 或更多）存在，它也是前生物合成的有用组分[23]。因此，前生物大气可能为大量小分子前体提供了来源，而这些小分子可能有助于有机合成和生命起源。

(4) 所有这些过程开始的准确时间非常重要，在地球历史的最初几亿年里，大气的组成几乎一定经历了演变。但在行星形成的过程中，由于撞击放气，CH_4、CO_2，可能还有 H_2 的初始含量应该更大。随着 CH_4 转化为 CO_2（也可能形成高级碳氢化合物），CO_2 转化为含碳岩石，所有这些气体的浓度在后来都应该降低了。在此期间，随着大气变薄，地表温度也可能降低，我们将在下一节回到这个问题。

(5) 在所有这些情况中，大撞击的发生应晚至 3.8 Ga。在 Nice 模型（专题 8.1）中，它们在这个时间附近集中发生；在其他模型中，它们分散在 4.4～3.8 Ga 之间。按“最大范围”，我们所说的撞击物大小上应能与产生了月球火山口 Imbrium 和 Orientale 的星体相类似，这两个火山口的直径均约为 1000km。要形成这样大小的火山口，需要的撞击物直径约为 100 km^2。而可能在 6500 万年前的白垩纪末期，使恐龙灭绝的撞击物，其直径可能为这个值的 10 倍，或质量为它的 1000 倍。一个直径 100 km 的撞击物，能够蒸发上层的 100 m 的海水，包括整个透光层[2]。因此，如果生命确实在 3.8 Ga 前出现，这是完全有可能的，那么它们就可能经历了多重全球灾难，这些灾难能够使近地表环境中所有有机生命体灭绝。在地下环境的有机生命体中，如洋中脊喷汽孔系统中的那些，可能在很大程度上没有受到这些事件的影响。正如在生命起源会议上讨论过多次的一样，这为核糖体 RNA 进化树根部附近，嗜热微生物（适宜生长温度高于 80℃的有机生命体）的主导地位提供了一个可能的解释[24]。

8.4 太古大气与环境

地球岩石的记录是从约 3.8 Ga 时真正开始的。目前为止我们已经讨论过在

那之前的时期，它被称为冥石代。在 3.8～2.2Ga 之间的时期被称为太古代。虽然比起较近的时期，我们对它知之甚少，但是在所有 7 块大陆上都有来自太古代的岩石，并且从这些岩石得到的太古大气和环境信息，亦有大量文献描述。在这简要的一章里，我们不着重概述这些证据。相反，我们着重几个方面，它们可以帮助读者理解我们对太古地球的已知信息，并引导我们思考没有岩石记录的更早的时代。

8.4.1 太古环境

我们考虑关于太古代和冥石代的基本环境问题，它们通常被称为微弱的年轻太阳的问题。我们认为在太阳系形成之时，太阳比现在要暗 30%，并从那时开始其亮度或多或少地呈线性增加（图 8.1）。虽然这是一种由星体进化模型提出的理论预测，但仍不失为一个很可能的结果（这种预测在最近 25 年内都没有发生大的变化，这一点可以支持这种假说）。这对地球环境的影响就很清晰了：如果在地球历史上其大气组成没有变化，在约 2Ga 前，全球平均表面温度将在冰点以下，这可由图 8.1 中上面的虚曲线看出。图 8.1 中下方的虚曲线代表了地球的有效辐射温度，可被认为是没有大气时的地表温度。两条曲线间的阴影部分代表了大气模型的温室效应，在这里假设为由目前水平的 N_2 和 CO_2 及不同量的水组成。在这次估计中，由于随地表温度升高，大气会含更多的水蒸气，因而温室效应随时间增强。

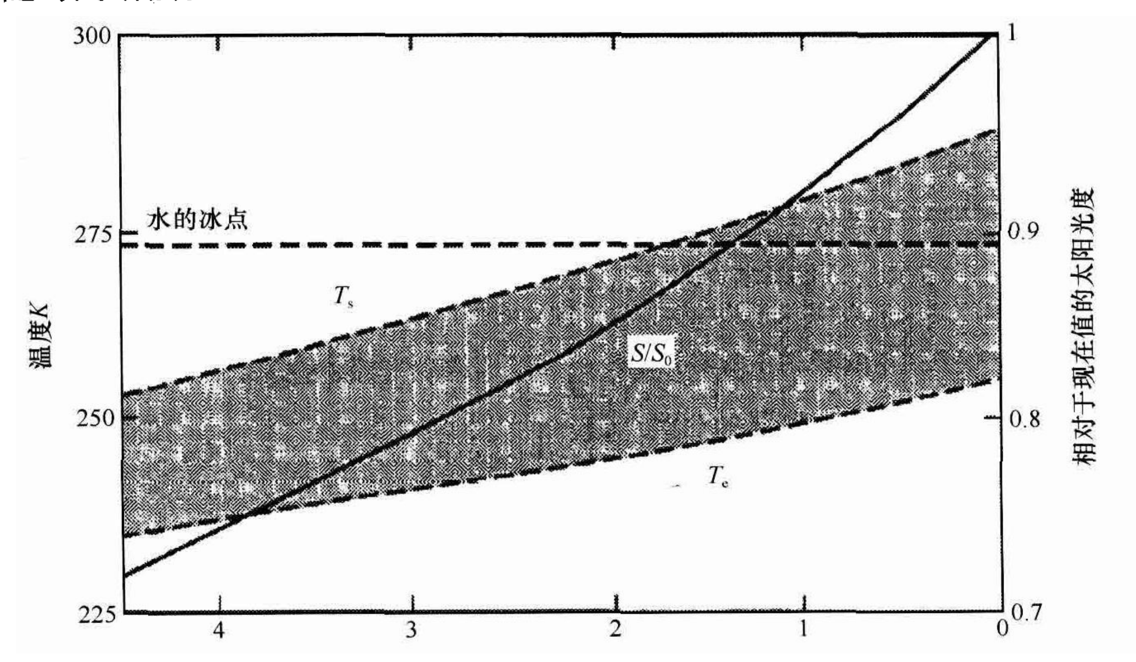


图 8.1　微弱的年轻太阳问题的示意图。实曲线为太阳相对于现在值的光度。下方的虚曲线为有效辐射温度，T_e。上方的虚曲线为平均表面温度，T_s。两条虚曲线间的阴影区域为温室效应的光度［根据 *Scientific American*，256 卷，Kasting[30]；版权（1988），得到 Scientific American 许可］。

对微弱的年轻太阳问题的可能解答已在其他地方回顾过了[3]，其答案几乎一定与过去温室气体的更高浓度相关。温室气体可吸收或发出 5～100 μm 的热红外辐射。最有可能的候选气体为 CO_2、CH_4，可能还有 C_2H_6（乙烷），后者最近才被提出，但它是一种强力的红外吸收物，可能提供 10℃或以上的温室效应。H_2O 也是一种强力的温室气体，但由于其饱和蒸汽压，它的浓度有限，更像是一种环境反馈，而不是环境的驱动力。

关于过去 CO_2 水平较高的论据如下：在很长时间内，大气 CO_2 浓度最初由无机碳循环或碳-硅循环控制。CO_2 通过火山注入大气，然后通过陆地上硅酸岩的风化移除，再在海洋中堆积成碳沉积物（图 8.2）。风化这个概念，是岩石被雨水中的溶解性物质侵蚀，发生相互作用而产生的物理和化学变化。硅酸盐的风化过程需要液态水，来达到一种较好的速率。因此，如果地球的平均表面温度降到水的冰点以下，如图 8.1 中提到的那样，硅酸盐的风化将会非常缓慢或完全停止，而来自火山的 CO_2 将在大气中积累。这种积累将会持续，直到多余的 CO_2 引起的温室效应足够高，将冰融化，使硅酸盐风化再次进行。计算显示[5]，CO_2 约为 0.3 bars 的分压就足以补偿太阳能流减少的 30%。作为比较，现在的大气 CO_2 浓度约为 300 ppmv 或 3×10^{-4} bars。因此，要补偿微弱的年轻太阳，需要的 CO_2 量约为目前大气水平（PAL）的 1000 倍。尽管这看起来很高，但它只是可利用的 CO_2 的很小一部分。地球大陆上的含碳岩石，含相当于 60～80 bars 的

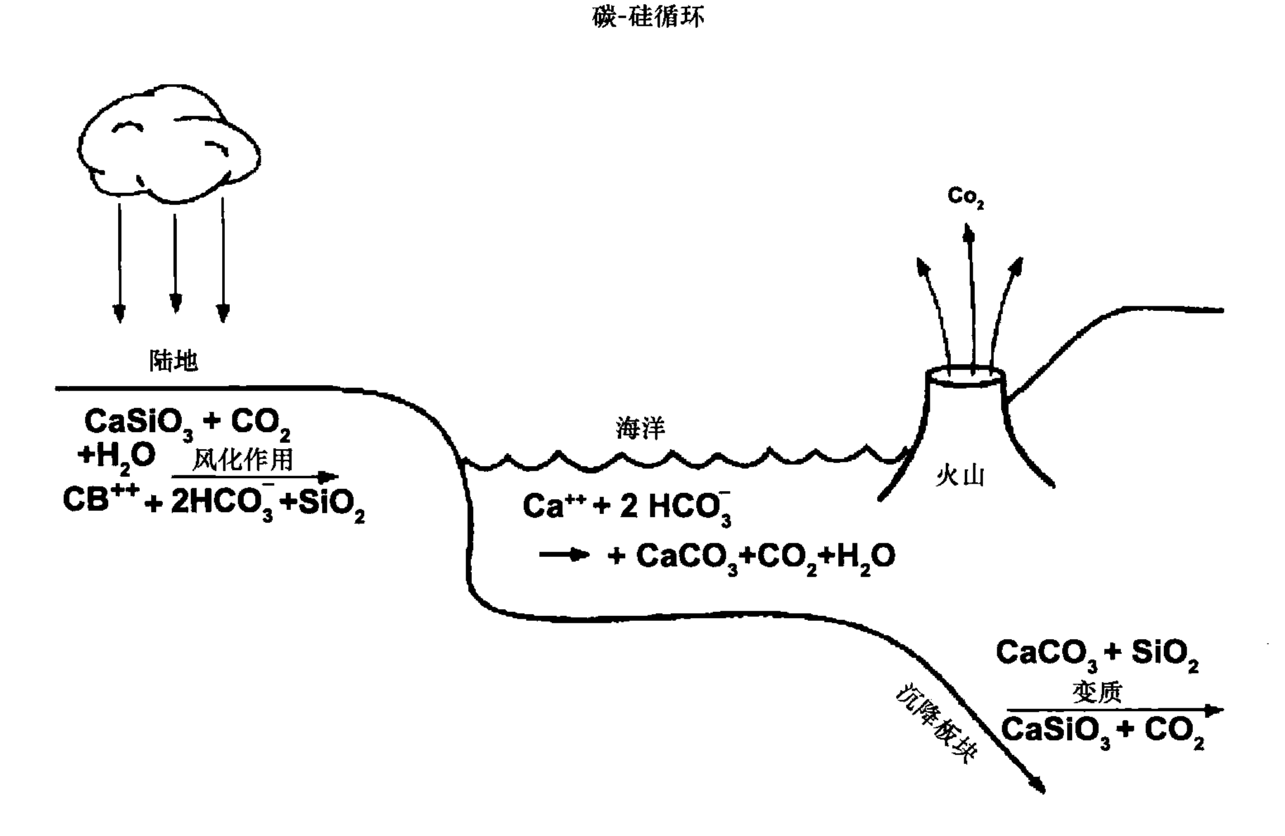


图 8.2　碳-硅循环图解。这种循环在很长的时间跨度内控制着大气 CO_2 浓度。

CO_2。只需要这些储备中的很小一部分，就可以补偿微弱的年轻太阳了。

此外，在早期地球，CO_2 可能是唯一的高浓度温室气体。在前面一节我们发现，CH_4 可能含量也很丰富。尤其在生命起源之后就更是这样，因为有机生命体可能产生大量甲烷。这些生物在现在被称为产甲烷菌（实际上高等植物也能产生一些甲烷，但这一点可以忽略，因为在太古代并没有高等植物）。产甲烷菌不是真正的细菌；其实，它们都只在 rRNA 树的太古区域分枝存在（见第 2 章）。多数生物学家都认为，它们进化得非常早，可能在整个太古时期都存在。那时大气中 O_2 水平很低（见下），导致 CH_4 的光化学寿命比现在至少长 1000 倍（10 000 年，而不是 10 年）。太古代 CH_4 的可能来源是生活在海洋和其沉积物中的产甲烷菌[22]，计算显示，这些 CH_4 的来源可与现今相比，因此，甲烷浓度不像今天的 1.6ppmv，而可能在 1000ppmv 以上，这样浓度的 CH_4 是一种有效的温室气体，可产生约 10℃的温室效应[25]。此外，在低氧环境中，甲烷的光解产生乙烷（C_2H_6），乙烷是一种更好的温室气体。因此，CO_2、CH_4 和 C_2H_6 的结合可能是维持太古地球温暖的机制。

其他关于太古环境的限制来自地球化学和地质学。一方面，黑硅石（SiO_2）中的氧同位素被用来证明太古环境是相当热的（约 70℃）[26]。实际上，来自碳酸盐的黑硅石数据和氧同位素数据显示，直到最近 4 亿年前，地球仍然相当温暖。另一方面，地质证据指出，地球在 2.9 Ga、2.4 Ga 和 0.6～0.7 Ga 都经历了冰川时期[27]，这些证据与早期温暖地球的氧同位素证据并不完全一致。以我们的观点，对这个问题的可能解答是，海洋中的氧同位素组成，随时间发生了变化。关于这个问题仍有很大争论，并且可能随时间变化。但是，我们认为，太古环境是温暖的，其他时期都不像氧同位素数据显示得那样热。

8.4.2 太古时期的 O_2 浓度

许多年来，地质学家和生物学家都很好奇，太古大气中存在着多少 O_2，以及大气中的 O_2 是从什么时候开始升高到了比较适宜的浓度。在 20 世纪 60 年代末期，地质学家 Preston Cloud 应用了各种地质学证据，包括红层、岩屑矿物和条带片含铁建造（BIF），显示在 2.0 Ga 左右，大气有一个富含 O_2 的时期。Cloud 的假说后来被 Heinrich Holland 扩展并详细阐述[5]。有如 Holland 所命名，大氧化事件（GOE）发生的时间，可以追溯到约 2.4 Ga，但 Cloud 的基本理念仍很可靠。多位地球化学家对这个假说提出了质疑，比较有代表性的是宾州州立大学的 Hiroshi Ohmoto，但他们的观点都没能得到支持。

约 7 年前，产生了一个具有重要意义的新进展，抓住了关于大气 O_2 含量升高的关键。James Farquhar 及其同事发表了一篇论文[28]，他们对古代岩石使用了多样的硫同位素，去验证 Cloud 的假说。硫同位素（^{32}S、^{33}S、^{34}S 和 ^{36}S）的一般

分馏或分离，通常取决于其相对质量（^{33}S 与 ^{32}S 差一个质量单位，而 ^{34}S 和 ^{32}S 差两个质量单位，因而，通常的分馏线性梯度约为 1/2）。在所有比 2.0 Ga 年轻的岩石中，硫同位素沿这样的质量分离线分布；但是，在比 2.4 Ga 古老的岩石中，硫同位素与之脱离（在 2.0～2.4Ga 的岩石与质量分离线接近，但不完全吻合）。作为他们研究中的一部分，Farquhar 等在实验室中，在没有 O_2 的条件下对二氧化硫（SO_2）进行光解，得到这个反应的产物具有质量独立的 S 同位素特征，即它们与质量分离线脱离。硫同位素数据显示出的事实与 Cloud 和 Holland 之前的地质学数据一致，这有力地支持了他们关于 O_2 演变的假说。但这并不意味着太古大气和冥石代大气中氧气含量为零。经如下反应，O_2 应持续不断地形成。

$$CO_2 + h\nu \rightarrow CO + O$$
$$H_2O + h\nu \rightarrow H + OH$$
$$O + OH \rightarrow O_2 + H$$

在这里，"$h\nu$"代表了一个能够分裂 CO_2 或 H_2O 的紫外光子。但是，由于它会产生光化学分解，O_2 不应在早期大气中积累。除了 CO_2 和 H_2O 这两种主要组成，火山气体也含少量的还原性气体，如 CO 和 H_2。如 8.1 节所述，James Walker 的论证显示，早期大气的 H_2 含量由火山口放气和氢气向空间的逃逸之间的平衡决定。这种大气氧化还原平衡的概念后来被提炼，应用于对水中氧化性和还原性物质的清除[3]。图 8.3 显示了一个包括了这种过程的典型光化学计算模型。O_2 通过光解在同温层中形成，在对流层（更低一层的大气）中通过与 H_2 反应而被消耗，这个反应由 H_2O 光解的副产物催化。这样的计算预测，近地表的 O_2 浓度应约为 10^{-13} atm。由于很有限，这样的浓度远远不足以对生物或前生物有机物合成产生影响。因此，模型和数据都显示出，大气 O_2 浓度在整个地球

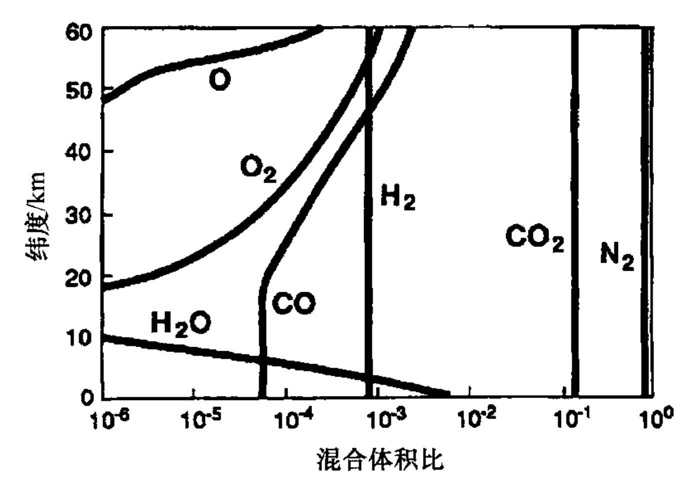


图 8.3 在早期还原性大气中一些主要气体的垂直分布［根据 *Science*，259 卷，Kasting[30]；版权 (1993)，得到 *Science* 许可］。

早期历史中基本都很低。

在 2.7 Ga 或更早的时期，演化出了产氧的蓝藻细菌[3]。它们出现的证据来自于保存在岩石上的复杂有机化合物——2-α 甲基荷烷，可能来源于蓝藻细菌的细胞壁或甾烷（甾醇的衍生物），推测后者来源于真核生物。真核生物，占据了 rRNA 树的 Eucharya 区域，是具有细胞核的生命体。真核生物中甾醇的生物合成需要自由氧，因此，在 2.7Ga 岩石上甾烷的存在意味着那个时候已经有氧气产生。一些早期蓝藻细菌可能共同生活在微生物垫，后来被岩石中薄的叠层石保存下来。另一些蓝藻细菌可能生活在海洋表面，尤其是上升流区域，那里富含营养物质，如混合的氮和磷。来自这些局部氧气绿洲的氧气团可能飘到不含氧的太古空气中，逐渐被消耗掉，就像还原性的污染物团逐渐被如今氧化性的大气消耗掉一样。这种"倒退"的大气一直存在，直到约 2.4Ga 时，O_2 占主要地位，H_2 和 CH_4 降到次要地位。

8.4.3 与冰河记录的联系

对以上论点最强有力的支持之一就是对冰河记录的成功解读。在前面一章里，我们提到在 2.9Ga 和约 2.4Ga 存在冰川，两者之中，2.4Ga 即古元古代的冰川提供了较好的档案，在至少三块大陆上，都发现了来自这个时期的冰川混积岩（或冰碛岩）。在北美洲，它们是南加拿大 Huronian 地层的一部分，紧邻休伦湖北岸。在 Huronian 有三块混积岩（图 8.4），分别命名为 Ramsey Lake（最底下）、Bruce（中间）和 Gowganda（最上层）。在 Ramsey Lake 混积岩下方，是 Matinenda 地层，在那发现了由岩屑形成的方铀矿（UO_2）和黄铁矿（FeS_2），这两种都是还原性的矿物，是低氧的证据（"由岩屑形成"的意思是，它们从母体岩石上风化下来，没有经过溶解过程。由于它们是还原性的矿物质，这只能在几乎没有 O_2 的条件下发生）。在 Gowganda 混积岩上方的 Lorraine 地层是一个成熟的红层（红层以赤铁矿的形式含有氧化铁，Fe_2O_3，这是高氧的证据）。因此，这一系列的冰川形成于同一时间，期间大气氧气水平首次升高。对于首先绘制了这个序列图谱的加拿大地质学家 Stuart Roscow 来说，这是一种巧合。但是，如果甲烷和乙烷在太古代对温室效应起了主要作用，而在原生代起的作用很小，那么当大气中 O_2 增加时

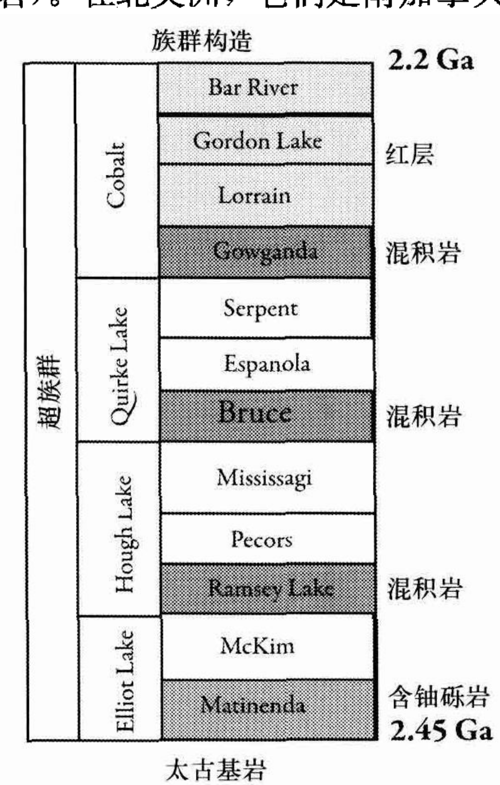


图 8.4 南加拿大 Huronian 超族群的地层顺序（见参考文献 31）。

环境变冷就一点也不令人吃惊了。

然而，这种看法有一些新的缺点。在 2.8～3.2Ga 的硫同位素数据比之前和之后的数据，质量独立分离（MIF）特征较小[29]，这可能是大气中 O_2 暂时增加的信号（如前所述），也有可能反映了大气组成的其他变化。如果在这个时期产生了有机物的霾层，或许能够同时解释低 MIF 值和 2.9 Ga 的冰川[27]，这种看法显然是纯理论的。在这里，我们仅仅指出了一些关于太古代大气和环境的现有模型，它们或多或少地与现存地质证据相一致，但这并不能意味着它们是正确的，当新的数据或模型出现时，它们可能会提供合理的解释。

8.5 结论

本章我们简单地综述了早期地球的大气和环境条件。敏锐的读者会注意到 8.2 节、8.3 节和 8.4 节之间的差异，在之前两节中我们描述了冥石代，而在后一节中描述了太古代，基于模型和数据的比较，后面的讨论得到了较合理的发展，而之前的两节猜测性较强，因为它们几乎全部都是基于模型的。我们关于地球形成和冥石代地球的模型都确实具有观测结果的支持，但它们所依赖的数据（如月球陨石坑记录）仅仅能间接体现早期地球的条件，并且，冥石代仅仅是生命起源的可能时期。这意味着对生命起源的探寻仍很艰难，因为只有少量关于早期地球条件的相关知识，这只是解决这个意义重大的有趣问题的又一个障碍。

致谢

感谢 Y. Watanabe 对图表所做出的帮助，以及 J. Siefert 对写这一章的鼓励。这项研究的资金由 NASA 的太空生物学和外太空生物学项目提供。

参考文献

扩展阅读

1. Walker JCG. Evolution of the atmosphere. New York: Macmillan 1977.
   • 虽然30年前出版，但仍提供了最好的有关氢逃逸与弱还原大气的讨论。
2. Sleep NH, Zahnle KJ, Kasting JF et al. Annihilation of ecosystems by large asteroid impacts on the early Earth. Nature 1989; 342:139-142.
   • 这篇文章提供了对重撞击的定量分析。要知道文章发表早于Nice模型的发展，因此假设撞击是连续性的。
3. Kasting JF, Catling D. Evolution of a habitable planet. Ann Rev Astron Astrophys 2003; 41:429-463.
   • 这篇综述讨论了行星可居性，微弱的年轻太阳，大气氧的出现，及大气的氧化与还原的平衡。
4. Kasting JF. Earth's early atmosphere. Science 1993; 259:920-926.
   • 该综述比文献3早，估计了需要多少CO_2才可以保持早期地球的温暖。
5. Holland HD. Early proterozoic atmospheric change. Early life on Earth. Vol Bengtson, S. New York: Columbia Univ Press 1994:237-244.
   • 该书综述了传统的有关大气氧出现的证据，当时文献28有关硫MIF的证据还未被发现。

专题文献

6. Oparin AI. The Origin of Life. New York: MacMillan 1938.
7. Miller SL, Urey HC. Organic compound synthesis on the primitive Earth. Science 1959; 130:245-251.
8. Rubey WW. Geological history of seawater. An attempt to state the problem. Geol Soc Am Bull 1951; 62:1111-1148.
9. Stribling R, Miller SL. Energy yields for hydrogen cyanide and formaldehyde syntheses: the HCN and amino acid concentrations in the primitive ocean. Origins of Life 1987; 17:261-273.
10. Pinto JP, Gladstone CR, Yung YL. Photochemical production of formaldehyde in the Earth's primitive atmosphere. Science 1980; 210:183-185.
11. Zahnle KJ. Photochemistry of methane and the formation of hydrocyanic acid (HCN) in the Earth's early atmosphere. J Geophys Res 1986; 91:2819-2834.
12. Halliday AN, Lee DC, Jacobsen SB et al. Tungsten isotopes, the timing of metal-silicate fractionation and the origin of the Earth and moon. Origin of the Earth and Moon. Tucson, AZ: Univ of Arizona Press 2000:45-62.

13. Matsui T, Abe Y. Evolution of an impact-induced atmosphere and magma ocean on the accreting Earth. Nature 1986; 319:303-305.
14. Genda H, Abe Y. Enhanced atmospheric loss on protoplanets at the giant impact phase in the presence of oceans. Nature 2005; 433:842-844.
15. Kress ME, McKay CP. Formation of methane in comet impacts: Implications for Earth, Mars and Titan. Icarus 2004; 168:475-483.
16. Schaefer L, Fegley JB. Outgassing of ordinary chondritic material and some of its implications for the chemistry of asteroids, planets and satellites. Icarus 2007; 186:462-483.
17. Ryder G. Bombardment of the hadean Earth: Wholesome or deleterious? Astrobiol 2003; 3:3-6.
18. Tsiganis K, Gomes R, Morbidelli A et al. Origin of the orbital architecture of the giant planets of the solar system. Nature 2005; 435:459-461.
19. Gomes R, Levison HF, Tsiganis K et al. Origin of the cataclysmic late heavy bombardment period of the terrestrial planets. Nature 2005; 435:466-469.
20. Kuhn WR, Atreya SK. Ammonia photolysis and the greenhouse effect in the primordial atmosphere of the Earth. Icarus 1979; 37:207-213.
21. Pavlov AA, Kasting JF, Brown LL. UV-shielding of NH_3 and O_2 by organic hazes in the archean atmosphere. J Geophys Res 2001; 106:23,267-223,287.
22. Kharecha P, Kasting JF, Siefert JL. A coupled atmosphere-ecosystem model of the early archean Earth. Geobiology 2005; 3:53-76.
23. Huber C, Wächtershauser G. Peptides by activation of amino acids with CO on (Ni,Fe) surfaces: implications for the origin of life. Science 1998; 281:670-672.
24. Gogarten Boekels M, Hilario E, Gogarten JP. The effects of heavy meteorite bombardment on the early evolution—the emergence of the three domains of life. Origin Life Evol Biosph 1995; 25:251-264.
25. Haqq Misra JD, Goldman SD, Kasting PJ et al. A methane/ethane greenhouse for the early Earth. Astrobiol; in press.
26. Knauth P, Lowe DR. High archean climatic temperature inferred from oxygen isotope geochemistry of cherts in the 3.5 Ga swaziland supergroup, South Africa. GSA Bull 2003; 115:566-580.
27. Kasting JF, Howard MT. Atmospheric composition and climate on the early Earth. Phil Trans Royal Soc Lond B 2006; 361:1733-1742.
28. Farquhar J, Savarino J, Jackson TL et al. Evidence of atmospheric sulfur in the martian regolith from sulphur isotopes in meteorites. Nature 2000; 404:50-52.
29. Ohmoto H, Watanabe Y, Ikemi H et al. Sulphur isotope evidence for an oxic archaean atmosphere. Nature 2006; 442:908-911.
30. Kasting JF, Toon OB, Pollack JB. How climate evolved on the terrestrial planets. Scientific Am 1988; 256:90-97.
31. Young GM. Stratigraphy, sedimentology and tectonic setting of the huronian supergroup. Toronto: Geological Assoc. Canada; Annual Meeting, Toronto, Field Trip B5, Guidebook 1991.

9 生物分子

王子晖

9.1 引言

遗传密码结构、大气氨基酸合成和陨石氨基酸这三项论据一同确定了最早活细胞的异养性质（见 1.4 节）。要产生一个异养的活细胞，需要从环境中获取一系列的生物分子，包括构成 RNA、蛋白质、细胞膜的结构单元，以及将它们拼在一起的能源，因此，前生物化学反应需要确定一个获取有机化合物的可行途径，是来源于地球上的前生物合成，还是来源于陨石或星际尘埃进行的地外传递。要判定一个有意义的前生物合成，标准如下所述[6]。

（1）当在合成位点有足够的起始物质存在时，才认为该前生物合成是可能进行的。

（2）反应条件必须与前生物大气和环境一致（见 8.3 节）。

（3）反应必须在有水或没有溶剂存在的条件下进行。

（4）反应的产量必须是显著的。

这些依据帮助我们确定哪些非生物反应模型是前生物的，哪些不是用现有知识能够解释的。实际上，有些时候乐观主义可能能够通过反推来证明。例如，众所周知，甲醛聚糖反应能够生成糖类，包括构成 RNA 的核糖，但核糖的产量通常很低，因而人们对甲醛聚糖反应对核糖的合成是否是“前生物”过程产生怀疑。但是，最近发现甲醛聚糖反应可以得到高产量的核糖，这增加了这一核糖可能来源的“前生物性”。

9.2 氨基酸

9.2.1 电火花合成

在斯坦利·米勒（Stanley Miller）历史性的实验中，通过加热产生水蒸气来模拟海洋的蒸发，接着混合甲烷、氨气和氢气，模拟水蒸气饱和的原始大气。当对混合物进行电击后，在水中发现了甘氨酸。一周之后，电击的烧瓶内部被油状物质覆盖，水变成棕黄色，分析显示其中含有不同种类的氨基酸和羧酸(表 9.1)[7,8]。

表 9.1 在甲烷-氨-氢气-水的混合物中得到的电火花产物[8]

	产量（mol×10^5）		产量（mol×10^5）
甘氨酸	63	蚁酸	233
丙氨酸	34	乙酸	15.2
肌氨酸	5	丙酸	12.6
β-丙氨酸	15	乙醇酸	56
α-氨基丁酸	5	乳酸	31
N-甲基丙氨酸	1	α-羟基丁酸	5
天冬氨酸	0.4	琥珀酸	3.8
谷氨酸	0.6	尿素	2
亚氨基二乙酸	5.5	甲脲	1.5
亚氨基乙丙酸	1.5	以上化合物总产量	15%

在电火花系统中 HCN 和醛类的存在，意味着产生 α-氨基酸的可能机制为 Strecker 合成（图 9.1）。即使早期地球大气所含有的还原性气体比最初电火花合成实验中使用得要少（见 8.3 节），氨基酸在大气内的合成仍能被以下发现支持[8~11]。

$$\underset{\text{乙醛}}{RCHO} + HCN + NH_3 \rightleftharpoons RCH(NH_2)CN \xrightarrow{H_2O} RCH(NH_2)\overset{O}{\overset{\|}{C}}-NH_2 \xrightarrow{H_2O} \underset{\text{氨基酸}}{RCH(NH_2)COOH}$$

图 9.1 α-氨基酸产物的 Strecker 合成[8]。

（1）除了放电合成，氨基酸也能通过其他物理途径产生，包括紫外线、高能射线和加热，这证实了早期地球上氨基酸非生物合成的基本可能性。

（2）对于蛋白质氨基酸和非蛋白质氨基酸，如 α-氨基丁酸和 β-丙氨酸，来自含碳 Murchison 陨石的氨基酸和电火花合成的氨基酸都是与之对应的。

（3）还原性更弱一些的大气也可以得到不错的氨基酸产量。

9.2.2 陨石

星际和陨石化学研究支持了在模拟地球前生物条件时，氨基酸非生物合成的可能性。人们以每年 4 个新分子的平均速率发现了许多星际间有机化合物（表 9.2）[12]，因此，有机合成在太空中极有可能是一个不断进行的过程，可通过如紫外光子等合成多种有机化合物。许多物质作为星际粒子从太空来到地球表面，它们以微米大小的彗星和星际尘埃存在。由于在早期太阳系中漂浮的碎屑更多，掉落的地外物质的量是今天的 100 万倍[13]。在太阳系增长的最后 10 亿年内，估计

地外物质下落带来的有机碳总量达到 $10^{16} \sim 10^{18}$ kg，这比现在生物圈中循环的有机碳总量还要大几个数量级[14]。

表 9.2 星际间分子。根据 Thaddeus[12]

原子数								
2	3	4	5	6	7	8	9	10 to 13
H_2	H_2O	NH_3	SiH_4	CH_3OH	CH_3CHO	$HCOOCH_3$	CH_3CH_2OH	CH_3COCH_3
OH	H_2S	H_3O^+	CH_4	NH_2CHO	CH_3NH_2	CH_2OHCHO	$(CH_3)_2O$	$CH_3(C\equiv C)_2CN$
SO	SO_2	H_2CO	$CHOOH$	CH_3CN	CH_3CCH	CH_3C_2CN	CH_3CH_2CN	$HOCH_2CH_2OH$
SO^+	HN_2^+	H_2CS	$HC\equiv CCN$	CH_3NC	CH_2CHCN	C_7H	$H{-}(C\equiv C)_3CN$	CH_2CH_2CHO
SiO	HNO	$HNCO$	CH_2NH	CH_3SH	HC_4CN	H_2C_6	$CH_3(C\equiv C)_2H$	$H(C\equiv C)_4CN$
SiS	SiH_2	$HNCS$	NH_2CN	C_5H	C_6H	HC_6H	C_8H	CH_3C_6H
NO	NH_2	$CCCN$	H_2CCO	HC_2CHO	c-CH_2OCH_2	CH_3CO_2H		c-C_6H_6
NS	H_3^+	HCO_2^+	CH_2	$CH_2{=}CH_2$	CH_2CHOH_2	H_2C_3HCN		$H(C\equiv C)_5CN$
HCl	NNO	$CCCH$	c-C_3H_2	H_2CCCC		CH_2CHCOH		
$NaCl$	HCO	c-$CCCH$	CH_2CN	HC_3NH*				
$AlCl$	OCS	$CCCS$	SiC_4	HC_4H				
AlF	CCH	$HCCH$	H_2CCC	C_5S				
PN	HCS^+	$HCNH^+$	$HCCNC$	C_4H_2				
SiN	c-$SiCC$	$HCCN$	$HNCCC$	HC_4N				
NH	CCO	H_2CN	H_3CO*	C-H_2C_3O				
SH	CCS	c-SiC_3						
HF	C_3	CH_3						
CN	$MgNC$	CH_2D^+						
CO	$NaCN$	$AINC$						
CS	CH_2							
C_2	$MgCN$							
SiC	HOC^+							
CP	HCN							
CO^+	HNC							
CH^+	CO_2							
CH	$SiCN$							
N_2	$AlCN$							
	$SiNC$							
	KCN							

陨落的彗星和星际尘埃布满地球表面，使早期地球的有机碳供应增加，但对初期的生命形式来说，这些碳相对难以被作为养分收集，来自陨石的有机碳则比较容易收集。对陨石样本的分子分析[15]得到了 10 种标准氨基酸，它们的遗传密码结构显示为 1 期氨基酸，这意味着在早期地球环境中，1 期氨基酸都是能够利用的[16,17]。这两种发现与大气氨基酸合成结果[10,11]完全一致，共同构成了生命异养起源的三个基础，而不是自养起源（见 1.4 节）。

这三个基础不仅显示前生物过程可以利用前生物环境中的 10 种 1 期氨基酸，即甘氨酸、丙氨酸、丝氨酸、天冬氨酸、谷氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸和苏氨酸，也显示了 10 种 2 期氨基酸的前生物不可利用性，即苯丙氨酸、酪氨酸、精氨酸、组氨酸、色氨酸、天冬酰胺、谷氨酰胺、赖氨酸、半胱氨酸和甲硫氨酸。事实上所有 1 期氨基酸都能通过弱还原性的 CO_2-N-H_2O 大气辐照获得[10,11]，这就完全满足了对大气氨基酸合成的需求。由于过度的不稳定性和不充分的合成，蛋白质最早并没有采用 2 期氨基酸作为结构单元。相反，它们是之后通过生物合成产生的（见 14.3 节）。环境中可利用的氨基酸尚且没有全部进入通用遗传密码，如 α-氨基丁酸和 β-丙氨酸（表 9.1），生命要么没有利用它们，要么就是尝试利用后又放弃了。对 β-丙氨酸，即使它最开始被利用了，最后被放弃也是不奇怪的，否则在全是 α-氨基酸的多肽中插入一个 β-氨基酸，将容易导致多肽链架构的破坏。

9.2.3 热液合成

与早期地球的弱还原性大气相比，从热液口喷出的富含 CO_2 和 H_2 的强还原性气体更利于有机合成。当以 3∶1 混合 CO_2∶H_2 气体，以及含氰化钾、氯化铵和甲醛的中性水溶液暴露于 150℃，10 个大气压下，模拟热液口的条件时，观察到了氨基酸的非生物合成，得到的氨基酸包括丝氨酸、甘氨酸、丙氨酸、天冬氨酸、异亮氨酸和谷氨酸，仍然都是 1 期氨基酸[18]。

9.2.4 手性

手性的 L-氨基酸组分对蛋白质结构和功能来说是必需的，而这种手性的起源长期以来都是前生物发展的挑战。在陨石氨基酸中，对异构体过量（异构体过量）的发现，指出了手性的地外起源（见 6.2 节、7.3 节）。碳酸钙是一种含量丰富的海中矿物质，它的选择性吸附，能够产生非外消旋的氨基酸。当碳酸钙吸附天冬氨酸时，被吸附的天冬氨酸 D/L 显示，异构体过量可以达到 10%。有趣的是，观察到的异构体过量值偏向 D 型或 L 型，取决于碳酸钙的表面晶形、异构体过量值越大，越有利于 L 型[19]。因而地球上的反应可能也促进了氨基酸的手性。

9.2.5 稳定性

在模拟前生物的条件下，即使某些有机物无法找到合适的非生物合成，进一步的实验仍可能会带来正面的结果。比较而言，若一种化合物极不稳定，导致它不能在前生物环境中存在，那么它的前生物不可利用性就更加难以克服，一些2期氨基酸就是这样的情况。在早期地球上，由于没有臭氧层，紫外线非常强烈。在以1.8mW/cm^{2}紫外线照射仅仅48h后，2期氨基酸中，色氨酸和组氨酸在pH=7的溶液中会有90%的降解，而苯丙氨酸和酪氨酸会有60%的降解[20]。在陆地或海洋中，通过岩石或深水对紫外线的遮蔽，这些氨基酸可能得到保护。但是，大气电火花合成的那些氨基酸，在被雨水冲下去之前，将难以逃脱紫外线的破坏。因此，即使有大量的非生物合成，最终的产量也将很低。

与紫外线破坏相比，热降解更容易发生。对热最敏感的氨基酸为2期氨基酸谷氨酰胺和天冬酰胺。在前生物环境中，任意一种化合物A的浓度可用非生物合成与降解的平衡来表示：

$$[A]\text{的变化率} = v_f - k_d[A]$$

在稳定状态下，变化率为零，因此，

$$[A] = v_f/k_d$$

式中，v_f表示形成速率；k_d表示第一降解速率常数。一种极端乐观的情况是，到达早期地球的所有波长小于260 nm的紫外光子都被用于非生物合成，这将使水圈中积累的氨基酸总量达到20 mol/L。即使所有氨基酸中的2.0%都是天冬氨酸，0.49%都是谷氨酸（即与电火花合成相对应），而且所有天冬氨酸和谷氨酸都分别来源于天冬酰胺和谷氨酰胺，那么在20℃，pH 7的条件下，通过v_f和k_d估计，天冬酰胺和谷氨酰胺的最大量仅为：

$$\text{最大 [Asn]} = 2.4\times10^{-8}\ \text{mol/L}$$

$$\text{最大 [Gln]} = 3.7\times10^{-12}\ \text{mol/L}$$

因此，由于其极低的热稳定性，Asn和Gln不可能积累到任何可被利用的程度[21]。所以，在前生物环境中并没有Asn和Gln。

9.3 RNA的组成成分

正如Orgel结论，前生物进化很可能经历了一个RNA世界：

“核糖体多肽合成反应由核糖体催化这个事实，使我们几乎能够确信，曾经有一个RNA世界。”[6]

因此，有大量研究都探讨有关RNA组分的前生物合成。

9.3.1 核苷碱基

Oro通过HCN对腺嘌呤的合成，开创了对核酸的前生物化学合成历程[22]。

氰化铵的回流形成了 HCN 的四聚体，再进一步多聚形成了一种深色的坚硬固体，从中可以重新获得腺嘌呤和鸟嘌呤[6]。要形成四聚体，需要 HCN 的浓度高于 10 mmol/L。由于在巨大的海洋中无法达到这样的浓度，而且对于 HCN 这样的易挥发分子来说，靠蒸发来浓缩也不可行，一个较有可能的方法是共晶凝固：若 HCN 的稀溶液降到低于 0℃，纯冰结晶出去，溶液变得越来越浓，直到达到共熔点－23.4℃，浓度为 74.5mol%，这时将产生大量腺嘌呤[23,24]，然而，也存在一些不依赖 HCN 的嘌呤合成途径[25]。嘌呤在 Murchison 陨石上的出现，亦证实了它们前生物合成的可行性（表 6.2）。

放电时，氮和甲烷混合物的主要产物是丙炔腈，它与两分子的氰反应会产生胞嘧啶，也能被水解为氰基乙醛，与尿素浓缩产生胞嘧啶（图 9.2）。胞嘧啶水解产生尿嘧啶[6,26,27]。由于胞嘧啶是不耐热的，即使在 0℃，其半衰期也只有 17 000年，因此需要持续补充，或者仅在较晚阶段为 RNA 或与 RNA 类似的生物多聚体利用。

$$N{\equiv}C-CH_2-CHO \quad + \quad H_2N-\overset{\overset{\displaystyle O}{\|}}{C}-NH_2 \longrightarrow$$

氰基乙醛　　尿素　　胞嘧啶

图 9.2　由氰基乙醛合成胞嘧啶。

9.3.2　核苷

在简单的矿物催化下，甲醛多聚形成糖混合物的反应，称为甲醛聚糖反应。它由 Butlerow 于 19 世纪发现，是一个独立循环自动催化的过程，可将甚至在星际尘埃（表 9.1）中都能找到的简单的 4 原子的甲醛，转化为复杂的糖混合物。由于甲醛聚糖反应产物的复杂，其中只有少量的核糖，因而有人对通过这种方式进行的核糖前生物合成提出了质疑。但是，在含有 Pb^{2+} 离子及催化量的戊糖途径中间体的情况下，产物中的戊糖包括超过 30% 的核糖，这提供了前生物核糖[30]。

碱性环境有利于甲醛聚糖反应，人们曾以为这种环境极少存在，但热碱系统可以提供这种环境，这与硼酸盐对核糖的稳定作用一起增加了早期地球上产生核糖的可能性[31]。而且，核糖比其他 5 碳或 6 碳糖更易渗入脂肪酸和磷脂膜，这可能是选择核糖作为 RNA 组分的一个重要因素[32]。含碳陨石 Murchison 和 Murray 包含各种多元醇，包括乙醇酸、甘油酸、二羟丙酮、丙三醇、赤藻糖醇、苏糖醇

和能够被转化成糖的核糖醇。这些陨石化合物，除了证明在前生物条件下非生物合成多元醇的可能性之外，也能为早期生命形式的核酸合成与中间代谢提供多元醇[33]。

当有氯化镁或海盐存在时，将D型核糖与次黄嘌呤共同加热，可生成达8%的β-D型肌苷。而在类似条件下，通过同样的反应并加上水解，D型核糖与腺嘌呤会生成3%的β-D型腺苷。还没有报道说通过核糖、尿嘧啶或胞嘧啶能够直接合成嘧啶核苷，但通过核糖或磷酸核糖与氨基氰和丙炔腈在水溶液中反应，可能得到α-胞嘧啶，它可以通过光转化成为β-胞嘧啶[6,34,35]。总体来说，在通向形成寡核苷酸的前生物反应中，核苷的形成是目前最薄弱的环节之一[6]。

9.3.3 核苷酸

在现代生物中，以ATP为磷酸供体，核苷被激酶磷酸化形成核苷酸。由于磷能够以适当强度形成多个化学键，加上磷酸具有三盐基性质，磷酸在能量转移过程中具有良好性能。无机多磷酸盐（PolyP），如环状三聚磷酸、线性多磷酸盐及一些特定焦磷酸盐，可在核苷酸合成中起到前生物磷酸供体的作用。由于在中性pH时，磷酸钙的溶解性很低，水圈中无机磷酸盐的浓度应低至0.1 nmol/L，即使如此，无机磷酸盐也可通过如下途径达到充足供应[36]。

(1) 虽然除了陨石上的陨磷铁矿外，许多普通含磷矿物都是微溶于水的，但估计在早期地球上，地壳含磷总量的5%都可能来自地外陨磷铁矿，通过陨石碰撞到达地球。Fe_3P是一种陨磷铁矿模型，可与水反应生成正磷酸盐和浓缩磷酸盐。

(2) 碱性的下降使得酸性钙盐透磷钙石（$CaHPO_4 \cdot 2H_2O$）析出，而高氨水和镁浓度利于鸟粪石（$MgNH_4PO_4 \cdot 6H_2O$）析出。透磷钙石和鸟粪石在加热时都能形成浓缩磷酸盐。

(3) 在强还原条件下，磷酸盐还原成亚磷酸盐。亚磷酸铵与核苷反应生成亚磷酸核苷，其二聚体易被氧化为磷酸二酯键。

(4) 焦磷酸盐和PolyP可在火山喷气口处获得[37]。

固态下与酸性磷酸盐共同加热，核苷可被转化为核苷酸，但这种转化效率不高[38]。在不含有机物时，将磷酸铵与尿素共同加热，会生成一种线性磷酸盐的混合物，而含核苷时，则会生成大量核苷酸。例如，当尿嘧啶与过量尿素及磷酸铵在100℃加热时，70%的尿嘧啶转化为尿苷酸的复杂混合物，也有可能调整反应条件，以利于生成特定形式的尿苷磷酸盐[6]。中性pH条件下，环状三聚磷酸与腺苷酸通过镁的催化产生9%的核苷酸，主要为2′，3′-cAMP[39]。干湿循环和Ni（Ⅱ）催化的介入使核苷酸产量增至30%，同时生成10%的2′，3′-cAMP和13%的ATP[40]。这些研究支持了前生物条件下PolyP的可用性，ATP作为高能

化合物，NTP 作为 RNA 聚合的底物。

9.3.4　非典型骨架

构造细胞和病毒 RNA 时都用到了 D 型核糖，这意味着对这种特定戊糖的利用由来已久。即便如此，也不能论断 RNA 之前没有其他含核苷碱基的生物多聚物。实际上，已经发现了一系列不同的含核苷碱基的生物多聚物，它们很可能是 RNA 世界出现之前复制子或基因的前体，先于用磷酸基团加在 3′-OH 及邻接呋喃核糖残基的 5′-OH 上的经典的骨架结构。

（1）由于核苷的非生物合成相对困难及对映异构体的交叉抑制，RNA 骨架中的核糖可能易于被非环式的类似物（如丙三醇、丙烯醛或赤藻糖醇的衍生物）取代下来，以克服这些不利影响[41]。

（2）吡喃糖 RNA（p-RNA），其主链上磷酸基团连接相邻吡喃核糖的 2′-OH 和 4′-OH，它的碱基配对系统比 RNA 牢固，选择性更强[42,43]。

（3）苏糖呋喃型核苷酸（TNA）的骨架中，5 碳的核糖被 4 碳的苏糖取代（图 9.3），可以形成同时含 DNA 和 RNA 的稳定异源双链[44]。TNA 的一个缺点是其骨架上没有游离羟基，这将降低其行使苏糖酶功能的能力，解释了前生物 TNA 被 RNA 彻底取代的原因。

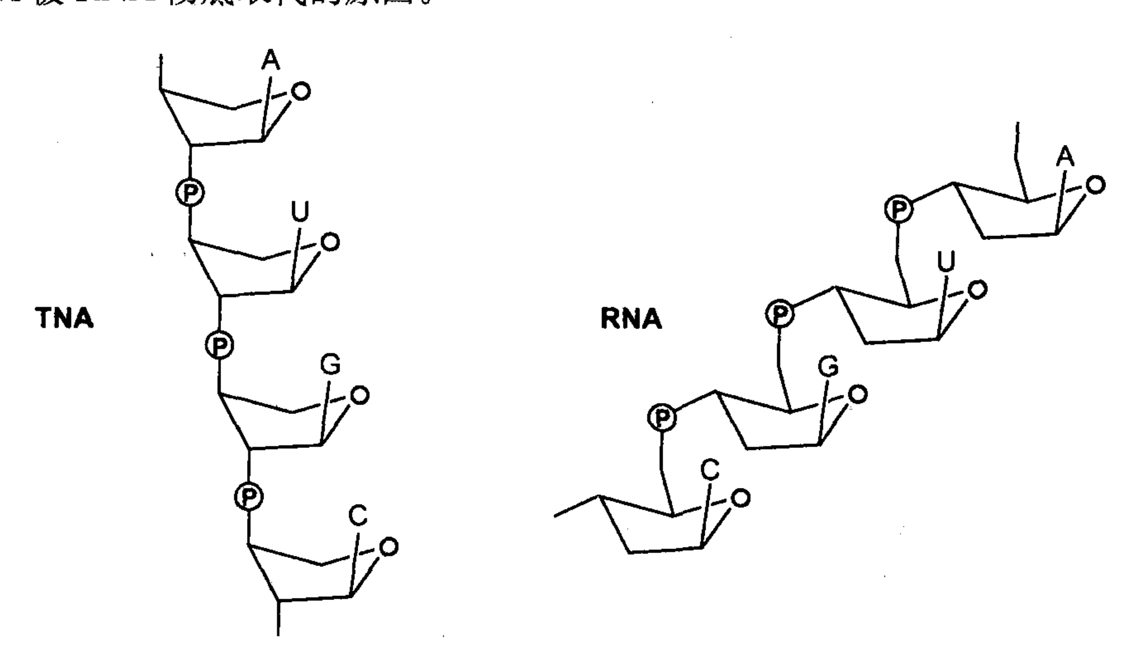


图 9.3　苏糖核苷酸（TNA）与核糖核苷酸（RNA）的比较。

（4）肽核酸（PNA）含常见的核苷碱基，与核酸配对，但不含磷酸盐[45]。PNA 被认为是一种前生物多聚物[46]，但目前还没有发现 PNA 单体的直接前生物合成[6]。而且，PNA 的整体灵活性比 RNA 低[47]，这意味着 PNA 在产生较小适配子或 PNA 酶方面都比 RNA 困难。

在前生物进化中，一个有价值的方向是，其他生物信息高聚物可能在RNA复制子之前出现，这些可选方案有助于对这个方向的研究。TNA的优势是可以形成TNA-RNA双链，可以完成从对TNA的早期利用或TNA和RNA的共同利用，到一个完全的RNA世界的转换，其中对复制子系统的破坏十分有限。

现在，生物体利用DNA基因，而病毒既利用RNA基因，也利用DNA基因。有证据表明，DNA的出现在RNA[48]甚至蛋白质[49]之后，如下所述。

（1）DNA的磷酸二酯骨架比RNA稳定得多。

（2）RNA聚合酶不具有校正功能，因而RNA基因组的突变率更高。

（3）胞嘧啶脱氨基形成尿嘧啶的过程，在DNA中会被检测并修复，而在RNA中却不能，因而RNA中的信息会降解，与DNA基因组相比，RNA基因组的突变率和出错率要高一些。

（4）紫外线照射对RNA产生的光化学影响大于DNA。

（5）脱氧核糖核苷酸是由核糖核酸还原酶对核糖核苷酸还原得到的。

（6）核糖核酸还原酶的催化机制中，对自由基的利用很可能只能通过一个复杂的蛋白质完成。

因此，基因的进化可能始于RNA、TNA或一些其他前体复制子，并发展成为RNA世界，最终，细胞中的RNA基因被DNA取代，而病毒可在DNA（如天花病毒）和RNA（如烟草花叶病毒）基因之间选择。早期复制子要么本身是生物催化剂，要么编码生物催化剂，这大大限制了最早的复制子的性质。RNA在行使核酶功能时符合其生物催化剂的要求（见11.2节）。蛋白质是高超的生物催化剂，但不能轻易自我复制。人们期待pRNA作为“p-核酶”、TNA作为“苏糖酶”、PNA作为“PN酶”行使生物催化能力，但仍不能确定它们在这方面的能力到底如何。

9.4 脂质

两亲性分子作为化合物，具有一个亲水头部，以及一个疏水尾部可和与水不互溶的有机溶剂亲和。类脂类物质为生物两亲性分子，带有长链脂肪酸或脂肪醇，其羧基或醇基为这个长链分子提供了亲水头部，而长的烷烃链提供了疏水尾部；或者在磷脂中，磷酸基团提供亲水头部，烷烃链构成疏水尾部。当被放入水中时，两亲分子趋向于形成微团，它们的尾部聚集到微团内部，远离其不亲和的水分子，而头部则朝向其亲和的周围的水分子。两亲分子可能也能自我装配成具有双分子层膜的囊泡，其中两亲分子排列成两层，其尾部相对，膜两侧的头部朝向水。

用有机溶剂提取Murchison陨石样本，并使之与水相作用，在其中观察到了自我装配的囊泡。这个令人惊奇的发现指出了地外两亲分子的存在，以及它们容

易形成膜囊泡的性质。两亲分子的前生物可利用性、自我装配成囊泡的性质，以及这些囊泡增大及分开的能力（见 12.4 节），促进了囊泡型生命形式的发展[50]。

在深海热液口及深层地下热含水层，也有可能产生两亲分子，那里的热能、矿物表面、还原性的氢及 CO_2 气体组分，能够促进合成各种有机物（见 1.4 节）。在一个模型中，利用草酸作为前生物反应的起始物质，在热液温度条件下进行水相 Fischer-Tropsch 反应，能够得到多于 5%的脂类，其中既有用于古菌脂类的正构醇，也有用于细菌及真核生物脂类的正构酸[51]。

9.5 能量来源

能量输入对前生物进化非常重要，这体现在两个水平上。第一，在前生物环境下进行的非生物合成，为异养的生命形式提供所需的生物分子，而这些非生物合成都是从自然界获取能量的，来源于闪电的电火花、紫外光子、雷产生的冲击波、宇宙射线、火山的热量及热液口等。第二，对生命或其在凝聚层中的前体，在囊泡内或吸附于矿物表面的有机物层来说，它们的吸能反应，如结构单元多聚作用，同样需要高能化合物的支持。

在现代生命中，高能化合物驱动着需能生物合成、逆浓度梯度转运、电导、机械收缩等过程，它们包括 ATP、其他高能磷酸盐及硫酯。这些能量载体通常是通过氧化还原反应交联或共价键断裂形成的。现在的生命体利用多种还原剂、氧化剂和能量底物。但是，对所有生命体来说，热动力基本规律和对高能键的需求都是最基本的，而且对前生物合成系统也会是一样的。从现代生命体得出的一个重要经验是，高能化合物极易相互供给能量，例如，肌肉中生成的能量较高的磷酸肌酸（−10.3kcal/mol），可以很容易地将它的磷酸基团转移给能量较低的 ATP（−7.3kcal/mol），而 ATP 可以与硫酯乙酰 CoA 进行交叉供能。关于高能化合物的前生物合成，人们已经提出了一些可能的机制，如下所述。

9.5.1 高能磷酸盐

热产生的 PolyP 可以为核苷三磷酸（NTP），包括 ATP 的合成供能，在前生物进化过程中的前囊泡时期，驱动 RNA 多聚及开放反应系统中的其他生物合成。现在，在所有已检测过的生命体中都发现了 PolyP，并已知其可在功能上替代 ATP[52]。

9.5.2 硫酯世界

通过底物水平磷酸化或膜连接的电子转运，可以生成高能磷酸键，而氧化还原反应对这个过程来说是必需的。在生命体中，它们释放的化学能通常被捕获并

储存在 ATP 中。但是，有人提出，在 ATP 出现之前，存在一个前生物的硫酯世界，其中硫酯是通用的高能化合物，驱动着代谢反应的发展[53]。根据这个理论，即使大气早期有氢流失，前生物铁离子也可以帮助产生一系列的还原化合物，包括难闻的硫化氢（H_2S），它可以生成有机硫醇（R-SH）。这些硫醇可以与羧酸（R′COOH）反应，通过三个途径产生高能乙酰硫酯（R′COSR）：

（1）在高温及酸性 pH 下反应；

（2）与将醛氧化成羧酸的反应交联，其中铁离子做氧化剂；

（3）与酮酸的氧化去羧基反应交联，其中铁离子做氧化剂。

生命体中最重要的硫醇化合物是辅酶 A（CoA），它的 β-半胱胺部分上含有一个硫醇基团。从生物学角度讲，由于 ATP 和乙酰 CoA 含有能量类似的高能键，它们可以彼此进行交叉供能，如通过乙酰 CoA 合成酶。从前生物来说，浓缩磷酸盐包括 PolyP 和硫酯也可彼此进行类似的交叉供能[53~57]。

9.5.3 Wood-Ljungdahl 途径

通过这个途径，产乙酸菌利用氢气做还原剂，厌氧产生乙酸来形成乙酰 CoA：

$$4H_2 + 2CO_2 + HSCoA \longrightarrow CH_3COSCoA + 3H_2O$$

在原核生物中，这个途径的核心酶包括一氧化碳脱氢酶。脱氢酶的机制可能是通过铁-镍-硫聚合物，利用 H_2 作为电子供体，使 CO_2 还原成 CO，这可能是由早期地热途径演化的酶途径的一种可能起源。在这个基础上，有人提出，对这种模式的乙酰 CoA 合成，有一种利用热液 H_2 的无机催化类似物，可代表最早的生化途径。在膜囊泡出现之前，这个途径可能存在于矿物微结构单元中[57]。

9.5.4 黄铁矿的形成

表面代谢理论假设，在还原性的火山喷气处附近，生命可能具有化学自养起源。最初的生命体及其原始代谢反应在带正电荷的矿物表面，如黄铁矿（FeS_2）附近经历了一段发展过程。带负电的 RNA 可以吸附在该表面并进行进化，这样可以克服原始浓度不够高的问题。为了进行碳固定，生命自养起源最关键的问题是自动催化循环是否可以与所需要的还原力共同作用，这种自动催化循环与还原的三羧酸循环相关，而还原力来自于黄铁矿的氧化形式[58,59]。

9.5.5 糖发酵

糖分解途径主要由葡萄糖到丙酮酸，再进一步形成乳酸、乙醇等。这个过程对所有三个领域的生命体来说，都是关键的供能途径。基于糖分解的顺序，一个葡萄糖分子转化成两个乳酸分子产生两个 ATP：

$$葡萄糖+2ADP+2Pi \longrightarrow 2乳酸+2ATP$$

由于在糖分解过程中，底物水平磷酸化产生能量的过程处在磷酸甘油醛和磷酸二羟丙酮的下游，丙糖模型认为这两种丙糖可能可以作为前生物能量底物[60]。更一般地说，不仅糖分解，其他利用大量的前生物化合物的发酵途径，如利用甘氨酸的途径，都可能起到了能量来源的作用[61]。

9.6 能源策略

9.6.1 优势和劣势

为了确定一个前生物进化的能源利用过程是否可行，我们有必要衡量一系列潜在能源的优势和劣势。

多磷酸盐

除去其钙盐的极低度水溶性外[36]，火山作用引起的 PolyP 的热形成[37]、由腺苷酸和三偏磷酸盐进行的 ATP 合成，以及磷酸盐的可用性，都造成了 PolyP 的优势。通过 PolyP 磷酸化得到的 NTP 可以多聚形成复制子，而且 PolyP 和 NTP 都能驱动代谢反应及生物合成[40]。不稳定性是 PolyP 的一个缺点[62]，它决定了 PolyP 形成的位置与其被利用的位置不能相隔太远。另一个缺点是，虽然现存的生命体具有许多膜转运体，能够从周围介质中输入溶质，但 PolyP 和核苷酸是不能被细胞直接摄入的[63]，我们没有理由认为前生命形式在这一点上能够做得更好。

硫酯

对于生命的异养起源，决定任何一种能量底物是否重要的关键，就是它的可利用性。在这一点上，Murchison 陨石中的可溶化合物（表 6.2）显示，陨石能够为前生物环境提供酮酸、醇酸和氨基酸。表 9.1 也显示，这些化合物能够在氢气-甲烷-氨水中通过电火花形成。同样地，氢甲烷氨水的电火花合成的最多产物为：丙氨酸、甘氨酸、α-丁胺酸，α-羟基-γ-丁胺酸、戊氨酸、肌氨酸和天冬氨酸[64]。现代的脱氢酶能够将醇酸转化成酮酸，例如，从乳酸到丙酮酸、苹果酸到草酰乙酸、异柠檬酸到 α-酮戊二酸。氨基酸氧化酶和转氨酶能够将氨基酸转化成酮酸，例如，以丙氨酸到丙酮酸、天冬氨酸到草酰乙酸、谷氨酸到 α-酮戊二酸。原始的无机催化剂也可能催化类似的转化反应，所得到的酮酸是早期地球上含量最丰富的能量底物之一[53]。通过原始的生物催化剂，这些酮酸可以产生乙酰硫酯。这些催化剂分别与丙酮酸脱氢酶，α-酮戊二酸脱氢酶和硫解酶类似，它们分别催化以下反应：

$$丙酮酸+CoA+NAD^{+} \longrightarrow 乙酰\text{-}CoA+CO_2+NADH$$

$$\alpha\text{-}酮戊二酸+CoA+NAD^{+} \longrightarrow 琥珀酰\text{-}CoA+CO_2+NADH$$

3 酮脂肪酸 CoA+CoA ⟶脂肪酸 CoA（少两个碳）+乙酰 CoA

因此，基于酮酸，醇酸和氨基酸在前生物环境中的可利用性，硫酯应该是一种出色的前生物高能化合物。

但是，没有复制子，代谢反应并不能自身完成高度进化，反之亦然，因此，一个完全没有 ATP，因而也就没有 RNA 或 RNA 类似复制子的硫酯环境，是不能够有效运转和进化的。考虑到这一点，在 PolyP 和 NTP 之前，仅仅优先利用硫酯的可能性就没有那么高了。

地热能

由于利用氢气做还原剂，Wood-Ljungdahl 途径适合于在地热地区的生命自养起源。同样地，黄铁矿的氧化形式可能能够产生还原力，通过硫代羧酸盐来驱动自养碳固定。虽然由于环境氨基酸的异养利用（见 1.4 节）和重要生物分子的热不稳定性[65,66]，有三方面的证据共同说明生命不是自养起源的，但是 Wood-Ljungdahl 和黄铁矿氧化形式途径都可以在生成生物分子作为营养物用于异养生命进化上扮演重要角色。此外，对生命的异养起源来说，要克服反应物的低浓度，对矿物表面或矿物微结构单元内容物的吸附具有相同的可能性。

糖类

现今，对于生命世界中具有重要意义的碳固定，光合作用是其主要途径。而糖类，作为光合作用的关键成分，是最重要的能量底物。生命起源前，可以通过甲醛聚糖反应由甲醛获得糖类，而 RNA 或 TNA 复制子的合成则需要核糖和苏糖的供应。但是，对糖类在原初代谢反应中起关键作用这个假设，以下观察结果是不支持的。

（1）糖类并不是最稳定的化合物。为了得到高产量，甲醛聚糖反应也需要碱性 pH。

（2）糖酵解是一个两段式的酶促反应。前半部分从葡萄糖到丙糖磷酸盐是糖途径，后半部分从 1，3-二磷酸甘油酸酯到丙酮酸为羧酸途径。代谢丙酮酸的三羧酸循环是一个类似的羧酸途径。在糖酵解后面的羧酸途径中，有两步底物水平磷酸化，但前半部分一步也没有。

（3）与前半部分反应相比，后半部分的酶高度保守，这意味着糖酵解途径最初是以糖异生从低到高的方向进行的，而不是糖酵解的从高到低[67]。起始于 RNA 世界，糖异生对将羧酸和氨基酸转化成 6-磷酸葡萄糖的过程来说是必需的，这个过程可以通过戊糖磷酸循环生成 5-磷酸核糖，再进一步生成 5-磷酸核糖-1-焦磷酸盐（PRPP），它是嘧啶和嘌呤核糖核苷的共同前体。

（4）通过膜化学渗透的氧化磷酸化产生 ATP 的过程，是利用来自于三羧酸循环和乙醛酸循环的羧酸作为底物的，而不是糖。

因此，在光合作用出现之前，糖类作为能量底物，其重要性不如羧酸和氨基酸。

9.6.2 PolyP-硫酯-ATP 策略

在明白了几种能源各自的优势与限制的前提下，我们有必要定义一个能源程序，使前生物生命形式在不同的发展阶段都能够得到不受中断的能量供应。一个三段式的 PolyP-硫酯-ATP 能源利用策略可能是一种合适的解决方案。

第一阶段：前囊泡

在初始时，PolyP 为 NTP 形成 RNA，TNA 和其他类似 RNA 的多聚物提供能量，而 PolyP 和 NTP 都能为生物合成和代谢供能。在 NTP 中，由于非生物合成的腺嘌呤产量高于鸟嘌呤，而相对于嘌呤核苷，嘧啶核苷的产量是供应瓶颈，所以 ATP 的含量比其他三种 NTP 都要高。这造就了 ATP 的前生物重要性，并使得现在的细胞都是利用 ATP 而不是其他 NTP 作为高能载体。NTP 多聚形成 TNA 或 RNA 复制子，它们在代谢物的指导下进化成了苏糖酶或核酶，扩大了代谢反应网，这些代谢产物包括来自于环境中的酮酸、羟基酸和氨基酸，而指导机制包含 REIM（代谢物引起的复制子诱导）和 REAS（稳定化引起的复制子扩增）所形成的苏糖酶或核酶，它们将酮酸、羟基酸和氨基酸转化成硫酯，以致 NTP 越来越依赖硫酯，而不是 PolyP 作为其合成的能量来源。

第二阶段：早期囊泡

当进化中的复制子-代谢系统被封入膜囊泡中后，囊泡内的生物进化就开始了。由于难以进入膜囊泡，外源的 PolyP 和 NTP 来源被阻断，因而在囊泡化的过程中，只有那些能够产生硫酯的前囊泡系统才能够存活。硫酯作为主要能量通货，符合硫酯世界[53]中硫酯取代 ATP 占主导地位的描述，但并不是说完全没有 ATP。硫酯为生成 NTP、合成 RNA 及复制子-代谢系统的继续进化提供能量。在这个早期囊泡阶段，膜仍远不够成熟，不足以支持高效的膜连磷酸化反应在最后一个全生命共同祖先，或 LUCA（见附录 15.1）的基因组中，没有细胞色素基因，这意味着膜是逐步进化的。由于没有与膜相关的 ATP 生成，在这个阶段硫酯在能量供应中占主导地位，这与现在的生物化学相符合。

(1) 现在硫酯只能激活羧基基团，而 ATP 可以磷酸化羧基、羟基、烯醇、酚和胍基。乙酰 CoA 和 ATP 中高能键的能量是类似的，而硫醇占主导地位的历史时期是在现代 ATP 主导的时期之前，这或许能够解释为什么乙酰 CoA 在生化途径中一直都是窄谱的，而 ATP 则通用的多。ATP 供能比限制较多的硫酯供能在功能上层次更高，这和人脑中新皮质的结构和功能比古老的神经结构层次更高类似。

(2) 对硫酯和ATP的平行利用是令人困惑的冗余现象，而这一点在多肽形成中尤为突出。在肽抗生素的合成中[68]，硫酯参与通过非核糖体肽合成酶(NRPS)进行的肽键形成过程，但不参与核糖体蛋白质合成。因为核糖体蛋白质出现较晚，并起始了酶代替核酶的过程，NRPS更像是一种肽合成的古老机制，以产生肽-核酶，它是介于核酶和酶之间的过渡生物催化剂（见14.2节）[69]。NRPS利用了硫酯，而核糖体蛋白质合成没有利用，这证明伴随着NRPS的发展，存在一个较早的硫酯时期；而伴随着核糖体蛋白质合成的发展，在硫酯时期之后形成了ATP时期。

(3) 甘油醛-3-磷酸脱氢酶能够利用糖类形成1，3-二磷酸甘油酸酯，它能够通过酶结合的硫酯中间物，将ADP磷酸化成ATP。糖酵解生成丙酮酸和脂肪酸氧化过程能够产生乙酰CoA，它与草酰乙酸共同浓缩生成的柠檬酸盐能够阻断生成ATP的三羧酸循环。在这些例子中，硫酯为生成ATP的重要途径开辟了道路。最初，当环境中的酮酸、醇酸和氨基酸几乎被异养生命挥霍殆尽，这些途径便能够有效地摄取糖类和脂肪酸作为能量底物，以供应ATP产生。这样，较早的硫酯供能便有效地为之后的ATP供能开展了新道路。

第三阶段：成熟的膜

在今天的生命体中，与膜相关的能量生成产生了大量的ATP，而过程中没有用上乙酰-CoA，这在原始生命形式中也是一样的。随着细胞膜结构和功能的成熟，利用质子梯度进行的膜连能量供应变得十分有效，由膜产生的高能键所占比例稳步上升，例如，在人体中，一个葡萄糖分子的完全氧化在细胞质中生成两个ATP，在线粒体膜上则生成28个ATP[70]。因而，成熟膜的出现和膜连能量供应共同确立了ATP作为高能化合物无可比拟的地位，它可以完全取代硫酯。ATP主导供能的这个最终阶段一直持续到了今天。

总之，在漫长的前生物/早期生物进化中，随着环境中羧酸和氨基酸的耗尽，进化中的系统面对着潜在的中断危险。封入了膜囊泡后，更发展了基于膜的能量产生方式。生命采取的能量策略必须能够对这些重大的变化产生有效的反应。PolyP-硫酯-ATP策略阐明了生命是如何在不同发展阶段调动不同能源，以应付这些潜在的断层。2000年间，人类的远距离地面运输经历了从人力到畜力再到石油驱动力的过程，而之后还可能向电力、生物能源和氢能源发展。因此，最佳的能量策略始终被相对可利用性、效率及能源代价的竞争力所支配，这对前生物、早期生命和人类来说都是一样的。

9.7 在地球的哪里?

通过研究活细胞中生物分子的非生物合成，人们合成了一系列结构单元，它们为前生物进化奠定了基础。可是，要使其中的不同合成方式达到最优，可能需

要相反的条件，例如，PolyP 和核苷的产生需要高温，而嘌呤合成和 RNA 的多聚则需要低温。由于生命的异养起源需要获取多样的环境养分，那么我们应该开始思考，哪里是早期生命起源的最佳地点，下面是几个可能的位置。

9.7.1　地表水

在以闪电供能的大气氨基酸合成中，生成的氨基酸可能会被雨水冲洗到水体表面，这形成了一个可能位置，如达尔文所说[71]："如果（嗯，仅仅是如果）我们认为有这么一个温暖的小水塘，其中含有各种氨和磷酸盐，还有光、热和电等，可以化学合成蛋白质，并进行更复杂的变化。今天，这些蛋白质会被生物迅速吞噬或吸收，但这在生命形成之前并不会发生"。在寻找温暖的小水塘这个问题上，达尔文事实上预言了太空生物学探索的重要目标——跟着水走就对了。

9.7.2　黏土和矿物质

各种各样的黏土是一类带负电的矿物质，包括高岭石（kaolinite）、蒙脱石（montmorillonite-smetite）、伊利石（illite）和亚氯酸盐，其中，蒙脱石被研究得最为广泛[72~75]。同样，也有许多带正电的矿物质，如双层氢氧化物（DLH）——绿锈或 Fe（Ⅱ）$_2$Fe（Ⅲ）（OH）$_6$。这两种类型的矿物质都分布广泛，并可以催化稀溶液中生物分子的化学转化和浓缩，从而在前生物化学反应中起重要作用。DLH 可以帮助形成并浓缩丁糖磷酸盐和戊糖磷酸盐，它们是 TNA 和 RNA 的原料[5,75]。蒙脱石催化形成含最多 50 个碱基的 RNA 低聚物，并引起顺序上、形态上和手性的选择，将统计上只有 0.2%的一些碱基异构体，积累到 13%，明显地增加低聚物的特异性[76]。前囊泡系统也有可能吸附在矿物质表面[58]或被封在矿物微室（micro-chamber）[57]中，以避免反应物的过度稀释，同时使不同前生命系统各自独立接受自然选择。

9.7.3　热液系统和撞击位置

第一个热液口是在探测海洋温度异常时，在 Galapagos Rift 附近发现的[77,78]。热液口是地壳上的裂隙，通常位于火山活动区，这里板块运动造成了岩浆的上升流，如中大西洋嵴（Mid Atlantic Ridge）和东太平洋海岭（East Pacific Rise）。海水在这些位置进入热液系统，并被火山加热到高达 400℃（图 9.4）。当这些过热液体流入周围冰凉的水中时，矿物质就会以每天高达 30cm 的速度沉淀出来并形成烟囱结构，它们可高达 60 m。那些能够喷射出富含硫的黑色矿物质的烟囱被称为黑色烟枪（black smoker）。相比较而言，白色烟枪温度较低，喷射的矿物质颜色较浅，富含钡、钙和硅。这些热液口是隐蔽在深处的热含水土层，能够进行一系列生物分子的非生物合成（见 1.4 节）。

除了深海热液口，陨石坑附近也会富含陨石带来的有机物（表 6.2）及能够产生磷酸盐的陨磷铁矿。由小行星和彗星产生的陨石坑也可能产生热液系统。撞击产生的冲击波能够生成氨基酸和其他有机物，而同时撞击物表面也可能带有宇宙中的大量物质[80]。

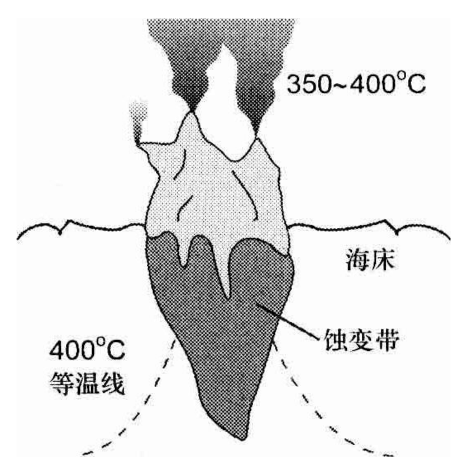


图 9.4 海床热液口（来自 Van Dover[79]）。

9.7.4 冰池

冰水混合物能够促进模板介导的 RNA 多聚化，它依赖于碱基配对和堆积，这个过程易在低温下进行。当存在比冰点略低的 Mg（Ⅱ）/Pb（Ⅱ）混合物时，将 90%以上的 4 种 RNA 单体大约等量反应，即可形成 5～17 聚物，并含有痕量的更长产物（见 10.4 节）。由于 RNA 及 RNA 类似物的关键作用，冰水混合物可能也是具有吸引力的位置[81]，这样的混合物同时也能促进嘌呤合成[23,24]。

9.7.5 混合型反应器

地表水、黏土、矿物质、热液系统和冰池都能提供重要的条件。那么哪里会是反应最快的地方呢？正如达尔文所说，速度是重要的，较晚产生的生命会被那些早来的吃掉。考虑到生命异养起源及迟早会需要 RNA，在有利于快速反应的因素中，能够提供多种营养源和适合 RNA 多聚作用的条件都会占着重要的地位。

这些考虑有利于一个假设的混合型反应器，其中聚集多种营养，适宜前生物进化[5]。图 9.5 阐释了这样一个反应器。在接近地球两极的火山区域中有一个具有潮汐的池塘，没有冷到完全结冰，但其低温足以连续产生或间接保持冰水混合物的存在。雨水混合着海水产生离子浓度梯度，适合于甲醛聚糖反应，并且足以溶解磷酸盐。周围间或有辅助性的黏土和矿物质，并毗连温泉、火山喷汽孔、陨石坑，甚至可能存在一个近岸的海中热液口。来源不同的有机物和浓缩磷酸盐进入这样的冰火反应器（Fire And Ice Reactor，FAIR），通过小溪、渗流和潮汐输送参与前生物进化。现今，在环太平洋火山带的北部可能会发现这样的冰火景观，如俄罗斯堪察加半岛、阿拉斯加各克水口，或北大西洋，如冰岛雷克雅未克由于其众多温泉而得名的冒烟湾（Bay of Smokes）。由于在早期地球上火山非常普遍，FAIR 也就会很常见。在数不尽的 FAIR 及其他众多混合或非混合型的反应器中，前生物装配过程会不断尝试发生，试图克服难以逾越的障碍。最终，一个在有或没有生命之间进行抉择的盘古竞赛（primordial bingo）在其中一个反应器中成功完成，这成功的故事至今仍在星际间回响。

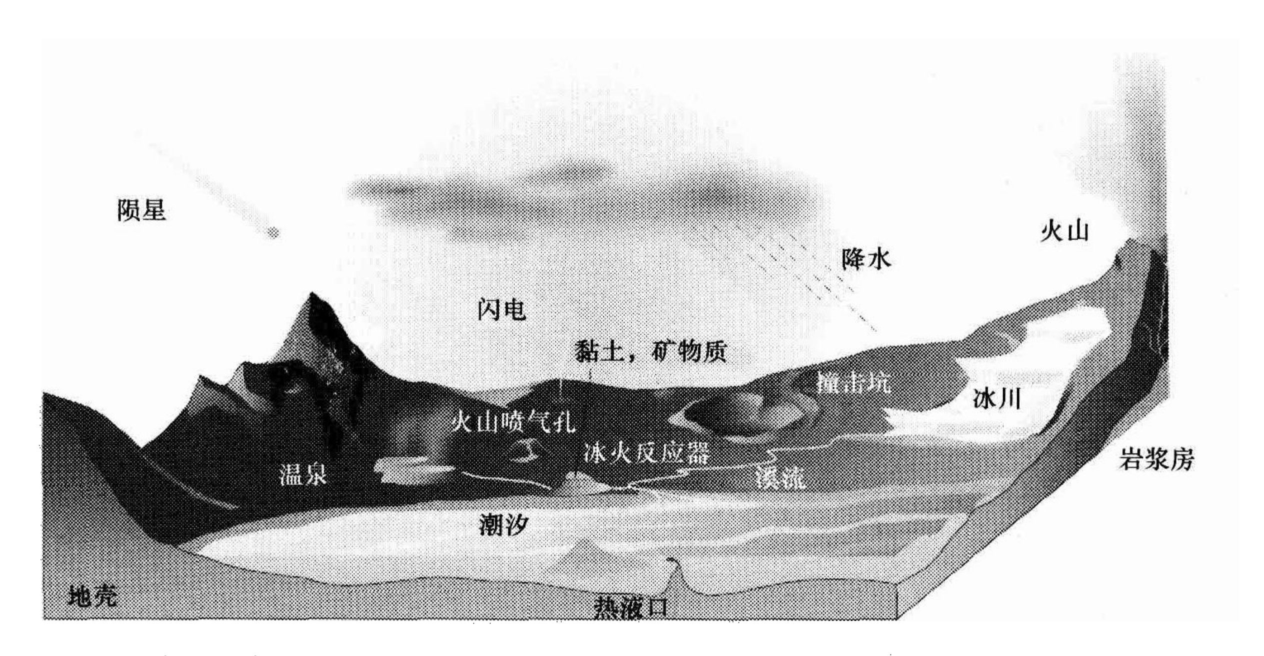


图 9.5　冰火反应器（FAIR），为 Mojzsis 等人从不同环境中收集生物分子的混合反应器之一，包括本例中的近海热液口[5]（见彩插）。

参考文献

扩展阅读

1. Brack A ed. The molecular origins of life. Cambridge University Press, 1998.
2. Deamer DW, Fleischaker GR. Origins of life. The central concepts. Jones and Bartlett, 1994.
3. Hazen RM. Genesis. The scientific quest for life's origin. Joseph Henry Press, Washington DC, 2005.
4. Jortner J. Conditions for the emergence of life on the Earth: summary and reflections. Phil Trans R Soc B 2006; 361:1877-1891.
5. Mojzsis SJ, Krishnamurthy R, Arrhenius G. Before RNA and after: geophysical and geochemical constraints on molecular evolution. In: Gesteland RF, Cech TR, Atkins JF, eds. The RNA World, 2nd ed. Cold Sprong Harbor Laboratory Press 1999:1-47.

专题文献

6. Orgel LE. Prebiotic chemistry and the origin of the RNA world. Critical Review Biochem Mol Biol 2006; 39:99-123.
7. Bada JL, Lazcano A. Prebiotic soup—revisiting the Miller experiment. Science 2003; 300:745-746.
8. Miller SL. The formation of organic compounds on the primitive Earth. Annals NY Acad Sci 1957; 69:260-275.
9. Miyakawa S, Yamanashi H, Kobayashi K et al. Preiotic synthesis from CO atmospheres: implications for the origins of life. Proc Natl Acad Sci USA 2002; 99:14628-14631.
10. Kobayashi K, Tsuchiya M, Oshima T et al. Abiotic synthesis of amino acids and imidazole by proton irradiation of simulated primitive Earth atmosphere. Orig Life Evol Biosph 1990; 20:99-109.
11. Kobayashi K, Kaneko T, Saito T et al. Amino acid formation in gas mixtures by high energy particle irradiation. Orig Life Evol Biosph 1998; 28:155-165.
12. Thaddeus P. The prebiotic molecules observed in the interstellar gas. Phil Trans R Soc B 2006; 361:1681-1687.
13. Bernstein M. Prebiotic materials from on and off the early Earth. Phil Trans R Soc B 2006; 361:1689-1702.
14. Deamer DW. Prebiotic amphiphilic compounds. In: Seckbach J, ed. Origins, Genesis, Evolution and Diversity of Life. Kluwer Academic Publishers 2004:77-89.
15. Pizzarello S. Meteorites and the chemistry that preceded life's origin. In: Wong JT, Lazcano A, eds. Prebiotic Evolution and Astrobiology. Austin: Landes Bioscience, 2008.
16. Wong JT. Coevolution of the genetic code and amino acid biosynthesis. Trends in Biochem Sci 1981; 6:33-35.
17. Wong JT. Coevolution theory of the genetic code at age thirty. BioEssays 2005; 27:416-425.
18. Hennet RJC, Holm NG, Engel MH. Abiotic synthesis of amino acids under hydrothermal conditions and the origin of life: a perpetual phenomenon? Naturwissen 1992; 79:361-365.
19. Hazen RM, Filley TR, Goodfriend GA. Selective adsorption of L- and D-amino acids on calcite: implications for biochemical homochirality. Proc Natl Acad Sci USA 2001; 98:5487-5490.
20. Wong JT. Evolution and mutation of the aminoacid code. In: Ricard J, Cornish-Bowden A, eds. Dynamics of Biochemical Systems. Plenum Press 1984:246-257.
21. Wong JT, Bronskill PM. Inadequacy of prebiotic synthesis as origin of proteinous amino acids. J Mol Evol 1979; 13:115-125.
22. Oro J. Mechanism of synthesis of adenine from hydrogen cyanide under possible primitive Earth conditions. Nature 1961; 191:1193-1194.
23. Sanchez R, Ferris JP, Orgel LE. Conditions for purine synthesis: did prebiotic synthesis occur at low temperatures? Science 1966; 153:72-73.
24. Schwartz AW, Joosten H, Voet AB. Prebiotic adenine synthesis via HCN oligomerization in ice. Biosystems 1982; 15:191-193.
25. Miyakawa S, Murasawa K, Kobayashi K et al. Abiotic synthesis of guanine with high temperature plasma. Orig Life Evol Biosph 2000; 30:557-566.
26. Ferris JP, Sanchez RA, Orgel LE. Studies in prebiotic synthesis. 3. Synthesis of pyrimidines from cyanoacetylene and cyanate. J Mol Biol 1968; 33:693-704.
27. Sanchez R, Ferris JP, Orgel LE. Cyanoacetylene in prebiotic synthesis. Science 1966; 154:784-785.
28. Levy M, Miller SL. The stability of the RNA bases: implications for the origin of life. Proc Natl Acad Sci USA 1998; 95:7933-7938.
29. Shapiro R. Prebiotic cytosine synthesis: a critical analysis and implications for the origin of life. Proc Natl Acad Sci USA 1999; 96:4396-4401.
30. Zubay G. Studies on the lead-catalysed synthesis of aldopentoses. Orig Life Evol Biosph 1998; 28:13-26.
31. Holm NG, Dumont M, Ivarson M et al. Alkaline fluid circulation in ultramafic rocks and formation of nucleotide constituents: a hypothesis. Geochem Trans 2006; 7:7 doi:10.1186/467-4866-7-7.

32. Sacerdote MG, Szostak JW. Semipermeable lipid bilayers exhibit diastereoselectivity favoring ribose. Proc Natl Acad Sci USA 2005; 102:6004-6008.
33. Cooper G, Kimmich N, Belisle W et al. Carbonaceous meteorites as a source of sugar-related organic compounds for the early Earth. Nature 2001; 414:879-883.
34. Sanchez RA, Orgel LE. Studies in prebiotic synthesis. V. Synthesis and photoanomerization of pyrimidine nucleosides. J Mol Biol 1970; 47:531-543.
35. Fuller WD, Sanchez RA, Orgel LE. Studies in prebiotic synthesis: VII. Solid state synthesis of purine nucleosides. J Mol Evol 1972; 1:249-257.
36. Schwartz AW. Phosphorus in prebiotic chemistry. Phil Trans R Soc B 2006; 361:1743-1749.
37. Yamagata Y, Watanabe H, Saitoh M et al. Volcanic production of polyphosphates and its relevance to prebiotic chemistry. Nature 1991; 352:516-519.
38. Beck A, Lohrmann R, Orgel LE. Phosphorylation with inorganic phosphates at moderate temperatures. Science 1967; 157:952.
39. Yamagata Y, Inoue H, Inomata K. Specific effect of magnesium ion on 2'3'-cyclic AMP synthesis from adenosine and trimetaphosphate in aqueous solution. Orig Life Evol Biosph 1995; 25:47-52.
40. Cheng C, Fan C, Wan R et al. Phosphorylation of adenosine with trimetaphosphate under simulated prebiotic conditions. Orig Life Evol Biosph 2002; 32:219-224.
41. Joyce GF, Schwartz AW, Miller SL et al. The case for an ancestral genetic system involving simple analogues of the nucleotides. Proc Natl Acad Sci USA 1987; 84:4398-4402.
42. Bolli M, Micura R, Eschenmoser A. Pyranosyl-RNA: chiroselective self-assembly of base sequences by ligative oligomerization of tetranucleotide-2'3'-cyclophosphates. Chem and Biol 1997; 4:309-320.
43. Eschenmoser A. Chemical etiology of nucleic acid structure. Science 1999; 284:2118-2124.
44. Schoning K, Scholz P, Guntha S et al. Chemical etiology of nucleic acid structure: the α-threofuranosyl-(3' → 2') oligonucleotide system. Science 2000; 290:1347-1351.
45. Egholm M, Buchardt O, Nielson PE et al. Oligonucleotide analogues with an achiral peptide backbone. J Am Chem Soc 1992; 114:1895-1897.
46. Miller SL. Peptide nucleic acids and prebiotic chemistry. Nat Struct Biol 1997; 4:167-169.
47. Sen S, Nilsson L. MD simulations of homomorphous PNA, DNA and RNA single strands: characterization and comparison of conformations and dynamics. J Am Chem Soc 2001; 123:7414-7422.
48. Lazcano A, Guerrero R, Margulis L et al. The evolutionary transition from RNA to DNA in early cells. J Mol Evol 1988; 27:283-290.
49. Freeland SJ, Knight RD, Landweber LF. Do proteins predate DNA? Science 1999; 286:690-692.
50. Deamer DW. *Prebiotic amphiphilic compounds. In:* Seckbach J, ed. Origins, Genesis, Evolution and Diversity of Life. Kluwer Academic Publishers 2004:77-89.
51. Rushdi AI, Simonet BRT. Lipid formation by aqueous Fischer-Tropsch type synthesis over a temperatrure range of 100-400°C. Orig Life Evol Biosph 2001; 31:103-118.
52. Rao NN, Liu S, Kornberg A. Inorganic polyphosphate in Escherichia coli: the phosphate regulon and the stringent response. J Bacteriol 1998; 180:2186-2193.
53. De Duve C. Blueprint for a cell. Neil Patterson Publishers. 1991:113-116.
54. Liu R, Orgel LE. Oxidative acylation using thioacids. Nature 1997; 389:52-54.
55. Weber AL. Formation of pyrophosphate, tripolyphosphate and phosphorylimidazole with the thioester N,S-diacetyl-cysteamine as the condensing agent. J Mol Evol 1981; 18:24-29.
56. De Zwart II, Meade SJ, Pratt AJ. Biomimetic phosphoryl transfer catalysed by iron(II)-mineral precipitates. Geochimica Cosmochimica Acta. 2004; 20:4093-4098.
57. Russell MJ, Martin W. The rocky roots of the acetyl-CoA pathway. Trends in Biochem Sci 2004; 29:358-363.
58. Wachterhauser G. Before enzymes and templates: theory of surface metabolism. Microbiol Rev 1988; 52:452-484.
59. Wachterhauser G. From volcanic origins of chemoautotrophic life to Bacteria, Archaea and Eukarya. Phil Trans R Soc B 2006; 361:1787-1808.
60. Weber AL. The triose model: glyceraldehyde as a source of energy and monomers for prebiotic condensation reactions. Orig Life Evol Biosph 1987; 17:107-119.
61. Lazcano A, Miller SL. On the origin of metabolic pathways. J Mol Evol 1999; 49:424-431.
62. Keefe AD, Miller SL. Are polyphosphates or phosphate esters prebiotic reagents? J Mol Evol 1995; 41:1432.
63. Carmeli C, Lifshitz Y. Nucleotide transport in Rhodobacter capsulatus. J Bacteriol 1989; 171:6521-6525.
64. Miller SL, Orgel LE. The origins of life on the Earth. Prentice Hall 1974:86-87.
65. Miller SL, Bada JL. Submarine hot springs and the origin of life. Nature 1988; 334:609-611.
66. Islas S, Velasco AM, Becerra A et al. Hyperthermophily and the origin and earliest evolution of life. Int Microbiol 2003; 6:87-94.
67. Ronimus RS, Morgan HW. Distribution and phylogenies of enzymes of *the Emden-Meyerhof-Parnas* pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism. Archaea 2003; 1:199-221.
68. Lautru S, Challis GL. Substrate recognition by nonribosomal peptide synthetase multi-enzymes. Microbiol 2004; 150:1629-1636.
69. Wong JT. Origin of genetically encoded protein synthesis: a model based on selection for RNA peptidation. Orig Life Evol Biosph 1991; 21:165-176.
70. Stryer L. Biochemistry. WH Freeman 1995; 552.
71. Brack A. The chemistry of life's origins. In: Seckbach J, ed. Origins, Genesis, Evolution and Diversity of Life. Kluwer Academic Publishers 2004:61-73.
72. Bernal JD. The physical basis of life. Routledge and Kagan Paul 1951;
73. Negron-Mendoza A, Ramos-Bernal S. The role of clays in the origin of life. In: Seckbach J, ed. Origins, Genesis, Evolution and Diversity of Life. Kluwer. Academic Publishers 2004:183-194.
74. Nikalje MD, Puhukan P, Sudalai A. Recent advances in clay-catalyzed transformations. Org Prep Proced 2000; 32:1-40.
75. Arrhenius GO. Crystals and life. Helvetica Chim Acta 2003; 86:1569-1586.
76. Ferris JP. Montmorillonite-catalysed formation of RNA oligomers: the possible role of catalysis in the origins of life. Phil Trans R Soc B 2006; 361:1777-1786.
77. Corliss JB, Dymond J, Gordon LI et al. Submarine thermal springs on the Galapagos Rift. Science 1979; 203:1073-1083.
78. *Baross JA, Hoffman SE. Submarine hydrothermal* vents and associated gradient environments as sites for the origin and evolution of life. Orig Life Evol Biosph 1985; 15:327-345.
79. Van Dover CL. The ecology of deep-sea hydrothermal vents. Princeton University Press 2000; 55.
80. Cockell CS. The origin and emergence of life under impact bombardment. Phil Trans R Soc B 2006; 361:1845-1856.
81. Monard P-A, Kanavarioti A, Deamer DW. Eutectic phase polymerization of activated ribonucleotide mixtures yields quasi-equimolar incorporation of purine and pyrimidine nucleobases. J Am Chem Soc 2003; 125:13734-13740.

10 RNA 世界的黎明：在可能的前生物条件下由单核糖核苷酸进行的 RNA 多聚化

Pierre-Alain Monnard

10.1 引言

由非生物合成的单核糖核苷酸进行 RNA 分子多聚化，是基于 RNA 双重功能（RNA 信息库和催化）的代谢作用初始的关键步骤，也是 RNA 世界假说的基本条件。这一章给出 RNA 非酶多聚作用方面研究的最新进展，并评估它们在探索 RNA 世界起始中的重要性。

10.2 前生物代谢中 RNA 是中心分子

现存的一般细胞生物服从人们已了解很多的分子生物学原理，依赖蛋白质进行核酸复制，并通过核酸来编码蛋白质。尽管现今复杂的细胞生物系统需要一个较简单的系统作为先驱，但对试图确定原始“生物化学”代谢性质的研究者来说，蛋白质与核酸的相互依赖一直都是个深奥的谜题（通常被描述为“鸡与蛋”的两难问题）。实际上，直到 1960 年 F. H. C. Crick[6] 和 L. E. Orgel[7] 提出 RNA 世界的概念，再由 W. Gilbert[8] 将其命名以前，这个问题的“满意”答案是不可想象的。这些作者假设有一种基于 RNA 的自主“生命体”或原始细胞，其中 RNA 链可以行使一些现在由蛋白质进行的功能，首先也是最重要的就是 RNA 复制，同时 RNA 也是遗传信息库。那时还没有发现催化型 RNA（所谓的核酶），这个想法也只是纯理论的，仅有的证据是普遍存在的辅酶可将核苷结合到自己的结构上。通过复制和突变进行的自然选择也是已知由简单生化系统进化到复杂系统的唯一机制。后来在现代细胞中发现了自我剪切的核酶，最近又证明核糖体多肽合成是一个核酶催化反应[9]，这增加了 RNA 世界作为现今 DNA-RNA-蛋白质世界前体这个假设的可信度。尽管从某些角度来说，要从了无生气的物质中产生细胞生命，RNA 世界似乎是不可避免的，但还不清楚在进化到生命的过程中它是不是第一个系统。实际上，现有的证据仍然不能解决起源及 RNA 世界“生物化学”的问题。

要实现功能化的 RNA 世界需要一系列事件（图 10.1），包括 RNA 单体的非生物合成及其到寡聚体的装配（可能有金属催化），以及寡聚体通过加

入单体或与另一个寡聚体结合而延长等。这些分子需要作为它们自我拷贝或复制的模板，而且所有这些在初始时都必须是非酶催化的。在这个阶段，一定存在过一个 RNA 池（大量 RNA 分子），并从中通过自然选择产生出一系列催化型 RNA（其中有依赖 RNA 的 RNA 复制分子），共同维持其在前生物环境中的指数增长。

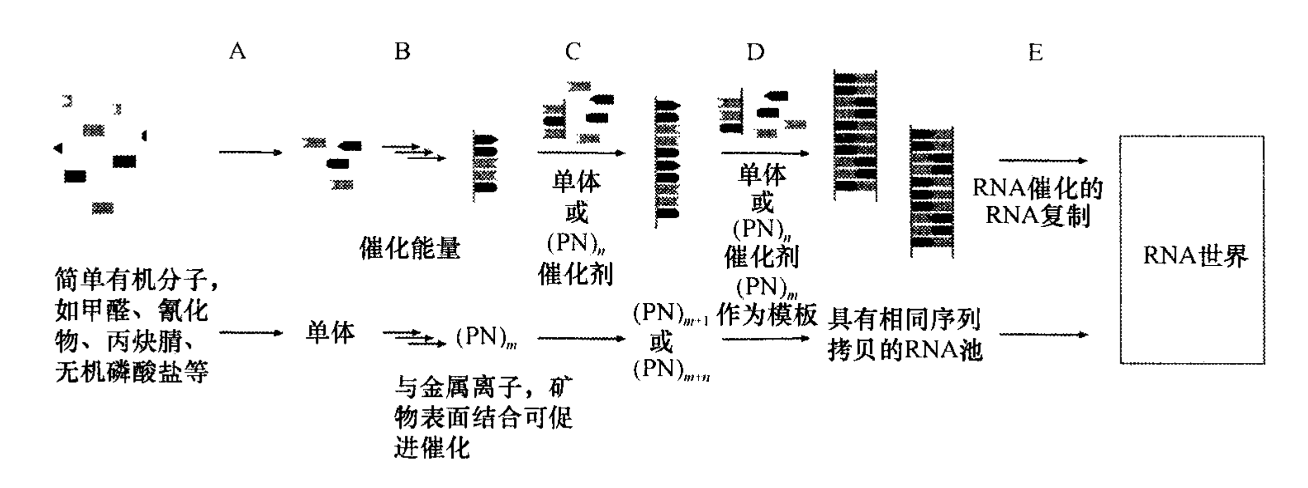


图 10.1　达到最初 RNA 世界的可能途径。N 代表任意核苷碱基，(pN) 为 RNA 寡聚体或多聚体。角标 m、$m+1$ 和 $m+n$ 代表多聚体的长度。A. 由简单有机物前体合成 RNA 单体。这一阶段依赖前生物化学反应。B. 单体的寡聚化作用。在可能有金属离子存在时，单体结合成寡聚体（多聚作用），而并没有蛋白质酶来催化这个反应。C. 合成较长的多聚物。多聚体需要一定的催化活性，可以来自于之前形成的寡聚体的结合，也可以来自于其他单体的添加。D. 非酶催化的模板介导的多聚化作用。要传递催化性 RNA 分子片段的序列，这种分子的非酶催化复制能力是必需的。这个过程不仅能增加催化活性，也能增大 RNA 反应池，同时，由于复制错误，也能产生新的催化剂。E. RNA 世界。当 RNA 反应池中出现这样一种分子，既能催化自身复制，又能催化其他功能性 RNA 分子复制，就形成了成熟的 RNA 世界。

10.2.1　RNA 多聚化中 RNA 活性的含义

催化过程需要催化剂识别靶分子。似乎早期核酶并不像现在的蛋白质酶一样具有底物特异性，这使得它们既能充分利用非生物合成的各种底物衍生物，又仍能保持相对较高的催化效率。正如蛋白质酶一样，核酶对靶分子的识别也是由 RNA 分子的构象（即它们的三维形状）决定的（图 10.2 及专题 10.1）。

活性 ⇒ 构象 ⇒ 二级结构模块 三级结构模块 ⇒ 核苷碱基序列

图 10.2　决定 RNA 活性的链结构。RNA 的催化活性取决于其构象，而构象来源于其二级结构和三级结构模块的相互作用。再进一步，这种模块的装配是由核苷碱基的序列决定的。

专题 10.1　RNA 结构的形成

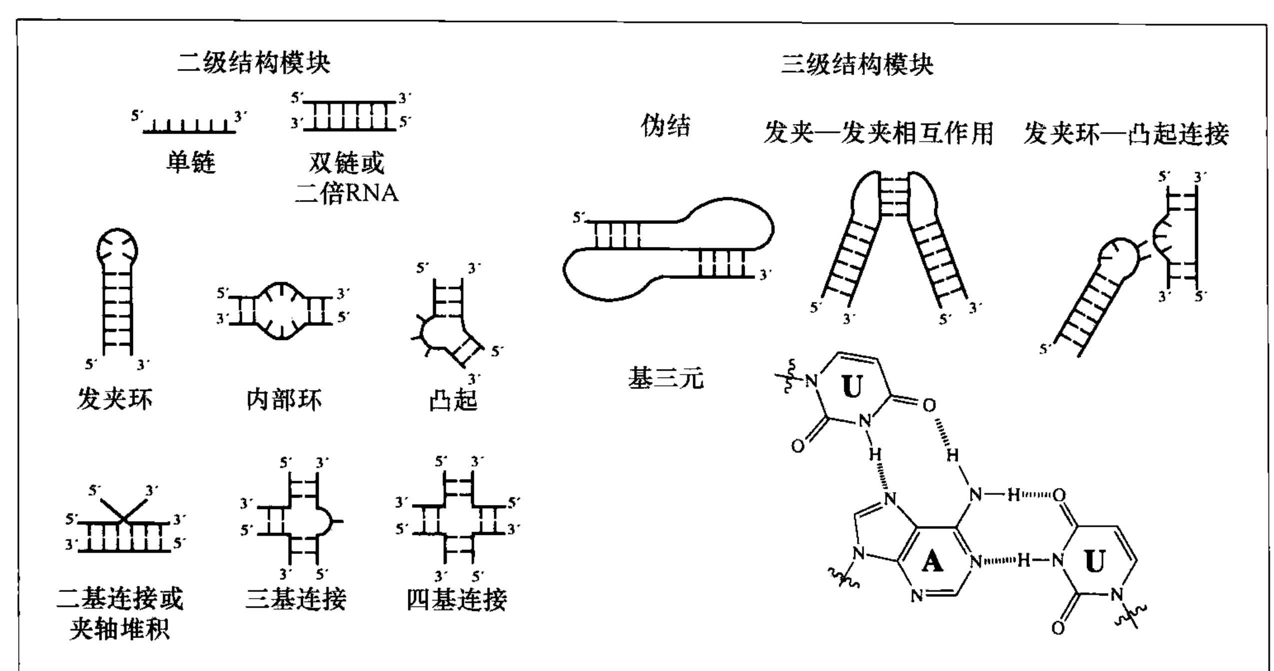
RNA 链能形成多种结构，共同完成分子在活体中的功能。与如今离体选择出的核酶一样，RNA 结构在 RNA 世界中可能也起重要作用。

结构主要取决于 RNA 分子的序列和它们的碱基配对方式，即两个碱基间形成氢键（H 键）的能力，以及非特异性碱基堆积的能力。其他特征，如沿磷酸核糖骨架的转角（专题 10.1.1 图 A)、碱基和核糖的构象（专题 10.1.1 图 B）也能通过限制分子构象而影响结构形成（见参考文献 1）。多种碱基对分为两类：普通的 Watson-Crick 碱基配对（专题 10.1.1 图 C）和非 Watson-Crick 碱基配对（专题 10.1.1 图 D)。

专题 10.1.1　影响 RNA 结构的因素。A. RNA 多聚物由 β-D-核苷酸通过磷酸二酯键连接形成。每个核苷酸由一个磷酸（深色）-核糖（浅色）骨架和一个碱基组成，一个杂环以 β 结构连接到核糖的 C1′上（碱基和磷酸都在核糖环平面之上）。碱基分为嘌呤[如腺嘌呤（A）和鸟嘌呤（G)] 和嘧啶［如胞嘧啶（C）和尿嘧啶（U)]。每个沿其骨架的单键（粗虚线）都能旋转。B. D 型核糖具有两种蹩起形式，C2′-在内或 C3′-在内。C. Watson-Crick 碱基配对：A · U 和 G · C 分别通过两个或三个氢键（虚线）相互作用。N-H 基团作为氢供体，而碱基环上氧的游离电子对（C=O）或环氮作受体。D. 非 Watson-Crick 碱基配对：左侧为“摇摆”碱基配对系统（G · U)；右侧为 RNA 中其他一些非 Watson-Crick 碱基配对方式。

通过两条链间的相互作用，或 RNA 上金属离子或蛋白质的结合形成的稳定结构属于二级结构和三级结构模块（专题 10.1.2 图)。

发夹为 RNA 折叠中最常见的模块。它们通常由较短的双螺旋主干和几个核苷酸构成的终端环组成。它们清楚显示了碱基序列和链构象间的关系，这种关系能够稳定 RNA 折叠，例如，环上序列为 UUCG 时，骨架和碱基的构象能在环内形成额外的摇摆配对（G · U)，并能使其余的环上碱基堆积起来，从而增加主干的热稳定性。



专题 10.1.2　二级和三级结构模块的例子。骨架由黑线表示，碱基为浅灰色的短线。两条相对的浅灰色短线代表碱基对。在 UAU 三碱基中，A・U 为 Watson-Crick 配对，而 U・A 不是。

RNA 模块的装配需要相对较长的多聚物：已知能够进行催化或分子识别的合成或天然 RNA 多聚体至少需要由 20～30 个单元组成，有较大的核苷碱基片段参与形成结构模块。离体选择出来的 RNA 连接酶和多聚酶则由更长的链构成，分别具有至少 50 个和 150 个单体。

这样的长度对 RNA 多聚化有一定的要求。实际上，一个指定活动若需要一个 50 残基（50mer）长度的 RNA 去完成，理论上将有 4^{50} 或 10^{30} 个不同的同长度 RNA 序列（设有 4 种不同碱基），这总量 3.5×10^{7} kg 的 RNA 并不能很快制造出来。即使这些序列中不止一个可能有活性，将它们减去后，浓度问题也一定不能忽视：对一个特定活动，活性分子的最低浓度都是十分必要的，在 RNA 世界初始时也必须考虑到，同时需要几个序列合作时尤其是这样。因此，RNA 多聚化过程需要生成一个 RNA 池，包含了大量催化活性分子（它们如何生成仍是个远未解决的问题）。

10.2.2　环境条件

不幸的是，关于原始地球的环境条件，仅有来自于生物圈的活动及板块构造运动的线索。不过，最近的研究表明，液态水和大陆早在 4.3×10^{9} 年前就存在了。也可以预期那时比现在具有较强的火山运动和较低的阳光照射水平，因此，产生生命的反应可能会在较宽范围的条件下进行，如温度、压力、离子强度、pH 等，并可能在陆地、地球深处、海洋中或这些区域的分界面进行。

10.3 RNA单体

尽管由前生物分子进行的不同RNA单体的合成超出了本章范围（见参考文献2），我们也有必要对这个问题的研究工作进行简述。我们有必要在此讨论一下，RNA世界从前生物有机混合物中出现的过程中，RNA单体合成（或缺失）的影响。

10.3.1 核糖核苷酸的合成

RNA单体——β-D-核糖核苷酸，是由三种不同类型分子——一个核苷碱基、一个核糖及一个磷酸基团形成的相当复杂的化合物（专题10.1），在装配成核糖核苷酸之前，其中每一个都需要作为中间产物进行合成（对磷酸基团来说，是使它能被利用）。

探索核苷酸前生物合成的实验通常关注这几个特殊部分之一。

- 展示了从大量非生物衍生物中合成所需嘌呤和嘧啶的途径。
- D型核糖可以由前生物物质甲醛通过甲醛聚糖反应产生，不过核糖只是所有糖类产物中相当少的一部分。而且，这些糖在pH=7时其分解半衰期从几小时到几年不等，这由温度决定。不过，与硼酸盐的络合作用及核糖衍生作用可以部分缓解这种情况。
- 将核糖与核苷碱基偶联能够产生核苷，这目前是核苷酸前生物合成中最薄弱的环节。不过最近有研究暗示了一种简化的合成，用磷酸盐衍生的核糖反应进行[10]。
- 核苷可以与无机磷酸盐反应生成5′-磷酸核苷酸。

尽管这些“具有前生物可能性的”合成利用了高纯度的反应物进行，但是它们中任意一环的产量都很低。而且像溶剂萃取这样的多重步骤也被用上，这显然不是前生物性的合成方式。最后，它们产生出的复杂混合物，其中有些是所需产物的异构体（具有相同化学式的分子，原子间通常也具有相同类型的键，但原子排布方式不同），它们很容易抑制多聚化或复制反应。

合成中的困难非常多，因此，原始地球上RNA最早是如何出现的这个问题应当谨慎考虑。举例如下。①核苷酸可能是由彗星和陨石带来的，它们的合成发生在完全不同于早期地球的条件下（温度、压力、pH）。如果考虑到每年地外天体带到地球上的碳化合物量（在地球刚形成后的几亿年间，超过10^{16} kg），这种说法可能是很合理的。② 可能还有其他更多的前生物分子参与形成RNA，它们是RNA世界之前，前RNA世界的组分。③RNA世界可能是通过催化反应网络形成的，它利用前生物有机分子去生成更复杂的分子。如果是这样，系统信息就与催化过程联系起来了，这种说法通常称为“代谢世界”。

前 RNA 世界的概念带来了对几种不同信息多聚物的研究，它们具有不同的骨架，其合成比较简单（图 10.3）。人们认为这个可能存在的 RNA 祖先应可以有效地将自身与 RNA 单体合成碱基对，方便两个世界中信息的转移。

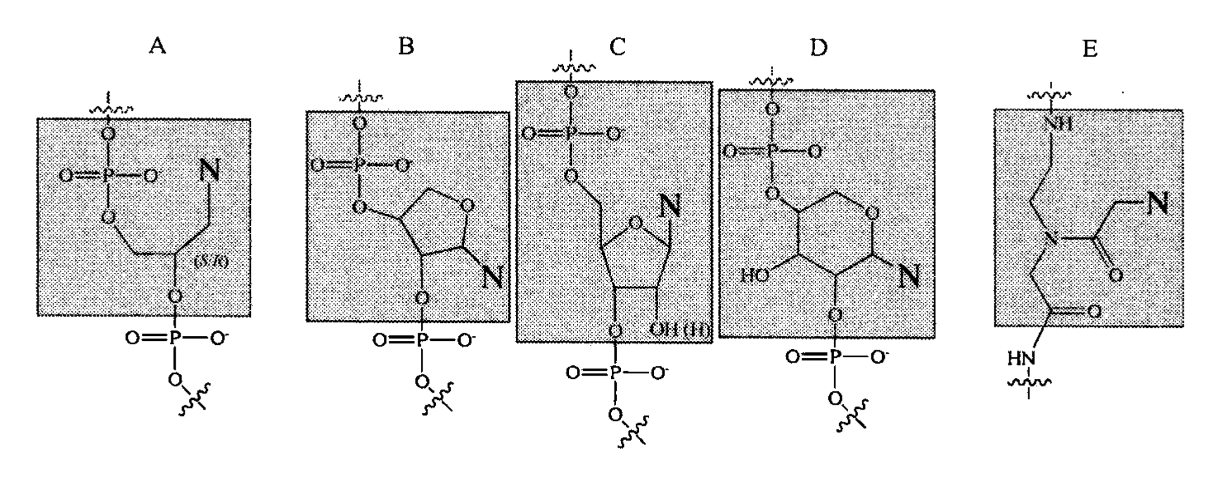


图 10.3 几种可选单体（深色标出）。A. 甘油核酸，（S/R）GNA；B. 苏糖核酸，TNA；C. 2′基团（OH 对 H）决定核糖或脱氧核糖核酸，即 RNA 或 DNA；D. 吡喃 RNA，pRNA；E. 多肽核酸，PNA。A 和 E 的骨架不是糖，B、D 和 E 可与其自身碱基配对。A、B 和 E 可与 C 配对，C 为中性核酸。

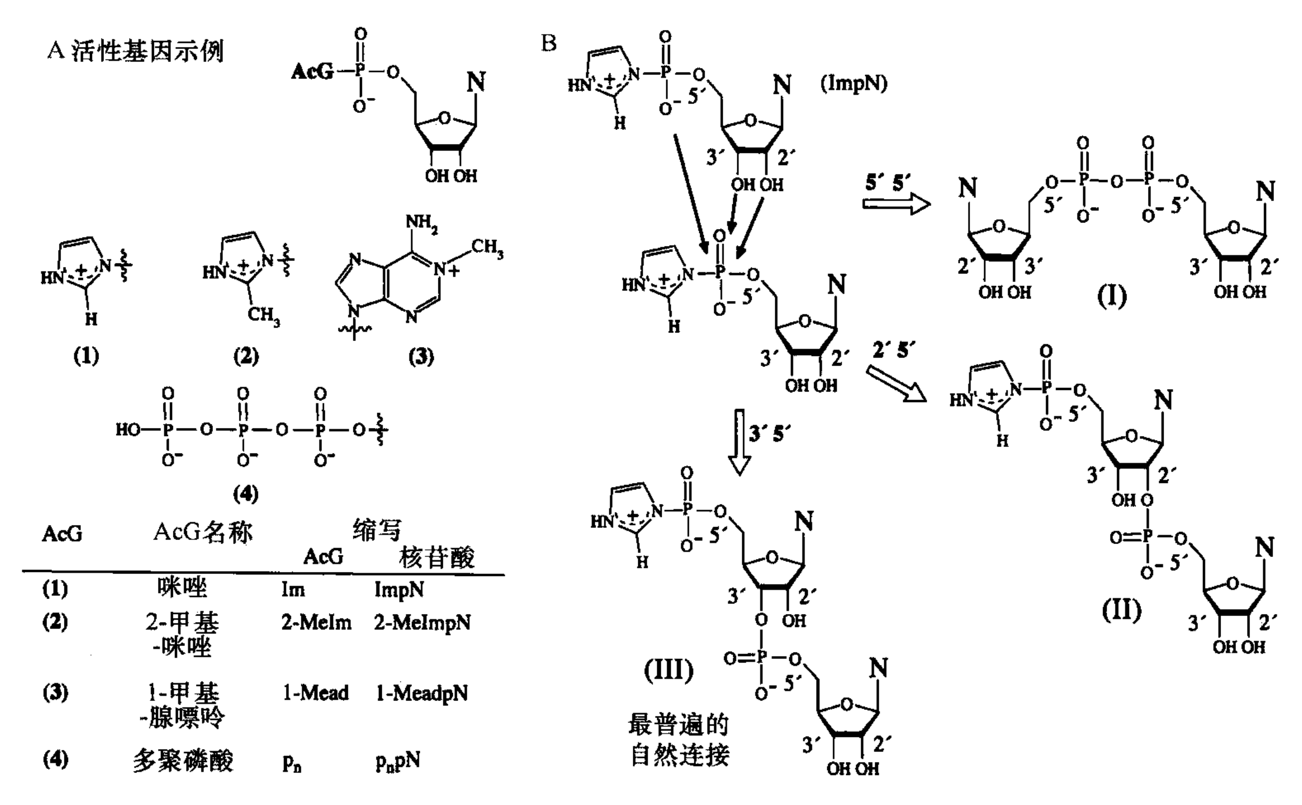
Eschenmoser 及其同事深入研究了糖环不同大小对核苷酸衍生的影响[11]，试图找出天然核酸选择核糖及脱氧核糖的原因。其他人则设计了新的多聚物，解决 RNA 合成或结构上的一些问题，例如，衍生出带有蛋白质[12]或甘油磷酸骨架[13]的核酸，这些衍生物十分有助于了解 RNA 的性质，尽管它们中的一些，如 GNA 和 PNA，具有前生物可能性，但其前生物合成，不论是不是非酶多聚化，都尚未有报道。还有，对这些过程实现的障碍有些已经得到了证明，如 PNA 单体会环化形成环形二聚体。

简言之，对 β-D-核糖核苷酸的初始合成还没有令人满意的方法，这可能也反映出我们对前生物条件知之甚少，还有可能恰好支持了 RNA 世界不是进化到细胞生命的第一步这个论点。但是，就算“代谢世界”这个假设是正确的，催化网络的信息在某些点也需要转化成基于信息或催化多聚物的形式。因此，要了解这个过程，对 RNA 非酶多聚化和复制的研究仍然是必要的。

10.3.2 RNA 单体的反应

为了保证 RNA 单体的多聚化，必须了解它们的反应。在水溶液中合成磷酸二酯键的标准自由能约为+5.3 kcal/mol，因此，多聚化不会自发进行，除非单体被化学激活了（即它们获得了储存于该单体及一个激活基团间化学键中的化学能），或得到外能（如热能）。要激活 RNA 单体，可以利用无机多磷酸盐（图 10.4），它们能促进多聚作用。但是，它们的多聚作用速率很慢，因此多数的

“前生物”多聚作用，一直都是通过将核苷酸激活为氨基磷酸酯进行的，通常为咪唑磷酸（图 10.4），即使这些分子在早期地球上可能并不存在。



AcG	AcG名称	缩写	
		AcG	核苷酸
(1)	咪唑	Im	ImpN
(2)	2-甲基-咪唑	2-MeIm	2-MeImpN
(3)	1-甲基-腺嘌呤	1-Mead	1-MeadpN
(4)	多聚磷酸	P_n	P_npN

图 10.4　活性 β-D-核糖核苷酸的反应。A. 活性单体。B. RNA 单体二聚化，区域选择性不同而产生的三种产物：（I）5′—5′或焦磷酸盐；（II）2′—5′；（III）3′—5′. N 代表任意碱基。

天然 RNA 通常具有 3′—5′连接，这是由多聚酶来保证的区域选择性（图 10.4）。由金属离子介导的非酶多聚作用倾向于产生具有所有可能连接的混合物，从而生成 RNA 类似物（在本章其余部分提到 RNA 时，即指的是 RNA 及其类似物的混合物）。它们的相对频率取决于其亲核羟基的相对反应性，也取决于金属催化类型和采用的介质。非酶合成寡聚体中连接的异质性亦可能阻止 RNA 的形式、功能及复制。

10.4　RNA 多聚化或 RNA 单体的浓缩

要具备“前生物可能性”，任何 RNA 多聚化过程都应当符合以下条件：

• 在水环境中进行，或至少能从水介质中得到反应物；

• 有少量复杂单体混合物存在，以作为浓缩的开始；

• 有金属离子催化剂存在，可以是游离离子或固体矿物表面；

• 聚合体的聚合速率必须高于解聚速率，这种动力学特征非常重要，因为由于熵效应，多聚反应在水环境中易被抑制；

• 反应必须能够选择性地合成具有高潜在活性的多聚物池。

现在我们来概述一下几种 RNA 单体多聚化的途径，也称为单体的自发浓缩，它们部分符合上述要求。

10.4.1 均匀水介质中的自发浓缩

在二价金属离子催化剂存在时（如镁、铅或铀酰离子），活性核苷酸应当能够形成延伸堆积（它使反应基团更为接近），这是由于核苷碱基具有相对疏水性，以及二价金属离子与带负电的磷酸及核糖羟基相互作用（图 10.5）。

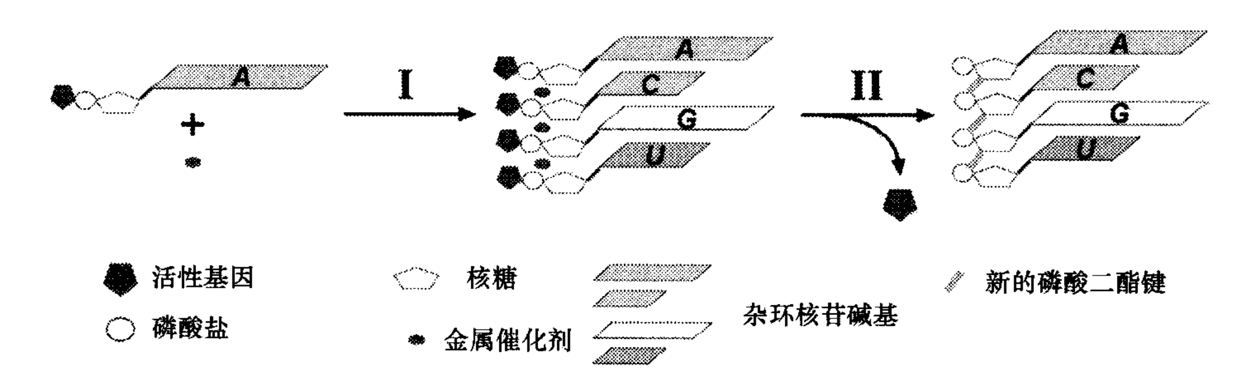


图 10.5 RNA 单体自发浓缩示意图。I. 堆（stack）的自我装配。II. 金属离子催化形成磷酸二酯键。

原则上，带有金属离子的活性基团应当能够克服能垒，从而使多聚化过程得以进行，但是，若催化剂和单体的浓度较低，初始产物就会是具有焦磷酸盐和2′—5′连接的二聚物。即使在如潮汐池这样的情况下（例如，随着水的蒸发，海滩上的小池子能够提供活性单体和金属离子催化剂的浓缩混合物），反应也只能产生少量产物，最多为四聚物[14]。这是因为活性物质的水解，以及亲水反应分子（尤其是 U）不能形成堆积。通过单体添加或短寡聚物结合进行的寡聚物延伸都不太有效。也就是说，在均匀水溶液中，非酶合成的寡聚物不能达到产物激活所需要的长度。

有了这些不成功的尝试，我们可以推断均匀水环境不能使反应物堆积达到有序。因此，需要为该系统引入一个特殊环境，能够与单体和催化剂相互作用并使之浓缩，从而减弱水分子活性，使它们达到有序。迄今已知有三类这种环境：矿物表面、冰水共熔相及脂双层结构。

10.4.2 在矿物表面，如蒙脱石上的自发浓缩

矿物表面的带电特性及常规结构可为 RNA 多聚化提供有利环境。不同的无机矿物质，如羟磷灰石 [$Ca_5(PO_4)_3OH$] 和蒙脱石（Montmorillonites）之类的黏土（图 10.6），可以促进 RNA 多聚化。经检测的几种普通黏土，多数都没有

表现或表现出很小的活性，但是，经化学处理的蒙脱土，弥散后称为同离子，具有显著活性[15]。

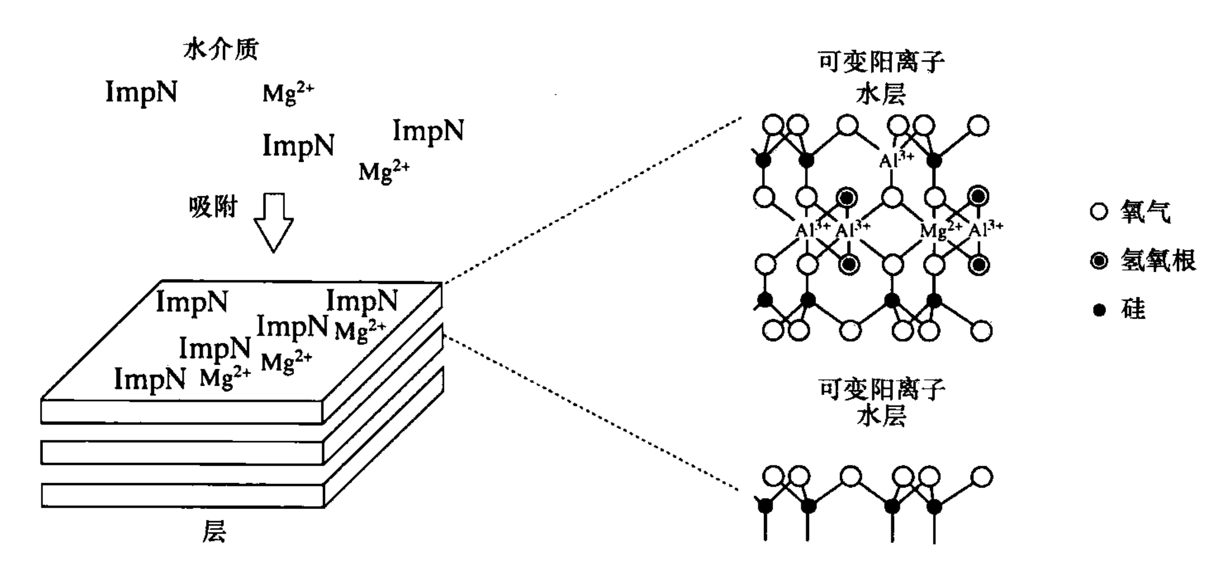


图 10.6　蒙脱石的结构及其对 RNA 多聚化作用的影响（见参考文献 16）。蒙脱石由两层硅四面体组成，层间由铝酸盐八面体连接。部分 Si^{4+} 和 Al^{3+} 被碱金属或碱土金属［如 Mg（II）］取代，或者是 Fe（III），这由侵蚀过程决定。当所有可变阳离子都被 Na^{+} 替换时，形成同离子蒙脱石。反应在表面进行，不在边缘发生。

活性核苷酸能够直接吸附在表面（通过疏水作用或带正电的活性基团及带负电的表面之间的静电作用）并（或）被表面及核苷酸磷酸酯之间的金属离子桥定位。吸附作用的程度取决于活性基团、矿物表面及水溶液的性质，而且，由于多重黏着作用，多聚体可以比独立单体吸附得更紧。

在含 $MgCl_2$（催化剂）的电解质溶液中，存在有小的引物寡聚物［短的脱氧核糖寡聚物，并加上 3′核糖核苷酸 $(pdA)_9pA$，或焦磷酸核苷酸二聚物 AppA］时，依赖于蒙脱石的 ImpA 能够自发浓缩生成长达 50 个单体单元的 RNA。在这个过程中，需要经常向单体堆积中补充活性单体（14d 一次），以避免单体水解带来的生长限制效应。若用大的杂环化合物作为单体的激活基团，如 1-甲基腺嘌呤，在没有引物的情况下也能够在仅仅 3d 内生成长度相似的寡聚物，且 4 种核苷酸都参与进来了[16]。

这种多聚化过程在产物核苷碱基序列和连接方式上（关于这些参数的实验测定，见专题 10.2）都具有选择性，例如，在 ImpA-ImpC 混合物中，所有可能二聚物都能形成，在可能随机合成的较长寡聚物异构体中，却只有少数能被检测到[17]。初生寡聚物的 3′端核苷碱基（嘌呤或嘧啶）及其区域特异性决定了延长反应。由于 3′—5′连接的末端核苷酸更易延长，所以在蒙托石上观察到的 2′—5′与 3′—5′连接的比例大大偏向于天然异构体（3′—5′），而均匀水介质中则以 2′—5′连接为主。最后，由于 3′端嘌呤本身活性比 3′端嘧啶高，所以偏向于形成富含

嘌呤的寡聚体。

专题 10.2 RNA 分析

非酶催多聚化形成的 RNA 链可通过不同方式按长度、碱基含量和连接方式进行分析。

凝胶电泳。通过与一个含已知长度 RNA 分子的 RNA 梯度进行比较，这种方法能确定 RNA 产物的长度（大于 5 聚物），分辨率可达一个碱基，且可以确定在引物上新增的碱基的连接方式。

色谱分析。高压液相色谱（HPLC）利用反相和离子交换柱进行 RNA 分析。反相色谱能够根据碱基含量和连接类型（2′—5′、3′—5′和焦磷酸盐）区分产物。这类方法可用于分析 RNA 多聚物的酶解产物，确定天然和非天然连接比例。还能用于确定在碱性条件或磷酸二酯酶催化下完全分解为核苷碱基后，RNA 中的碱基含量。离子交换层析用于确定较短寡聚物的长度和它们的空间构型。

酶消化。核酸酶或 RNase 能特异性裂解 3′—5′连接，对裂解位点两侧的碱基和 RNA 结构（单链或双链）都具有序列特异性。这些酶和凝胶电泳与色谱分析一起，能够帮助我们确定天然和非天然连接的比例。

质谱分析（MS）。尤其 MALDI-TOF MS 可用于较短寡聚物的序列分析。这种方法使人们能在不损伤生物多聚物的同时，分析出它们的重量。质量分辨率可低至一个质子，进行相对精细的序列测定。

测序。这种方法能确定天然 RNA 和 DNA 的碱基序列。现今，已能用酶催化，将 RNA 片段反转录为 cDNA 或完整 DNA。再将这些 cDNA 扩增，插入质粒，转染到细菌中进行进一步扩增，再对 DNA 质粒产物进行测序。非酶催合成的 RNA 为混合连接，会使反转录极为困难。

10.4.3 冰水共熔相

近年来对行星体的探索清楚地证明，在地球以外的宇宙其他位置，水分子经常表现为冰。在早期地球上，太阳辐射强度比现在弱（约 30%），因此可能会有冰水混合物。

冰水系统是 RNA 单体自发浓缩的有趣环境，这是因为这种系统能够非常有效地浓缩溶质（图 10.7），从而使它们组织起来。冰冻可以通过给样品脱水来减小系统中水的活性，而低温则能够保护 RNA 不被热解。相关的实际过程（简单浓缩效应与冰表面的相互作用等）还没有得到阐明，不过只有当冰核形成时多聚化才能发生。过冷溶液并不利于高效反应。

若混合物包含所有 4 种咪唑激活的单核糖核苷酸，同时存在金属离子，尤其是 Mg（II)/Pb（II）混合物，以准等摩尔（quasi-equimolar）在−18.4℃（比冰点略低的温度）孵育超过 36 天，观测到所有单体几乎都（达 90%）形成了中等长度的混合寡聚物（达 15～30 聚物）[18]。与黏土相反，这种反应是在活性基团与

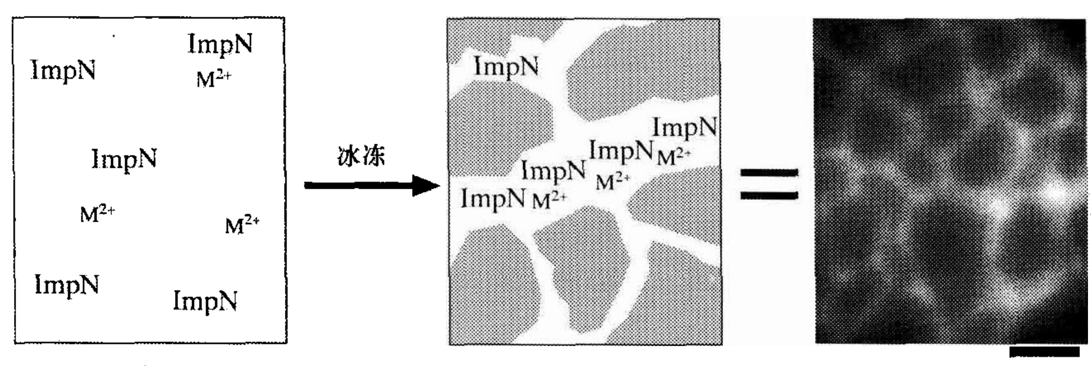


图 10.7　冰水共熔相的形成。活性单体（ImpN）和金属离子催化剂（M^{2+}）的稀溶液在其冰点之下，共熔点之上凝固（即在此温度范围内整个样品都凝结为固体)。凝固过程中，溶质在纯冰晶体（见含荧光染料的混合反应物的显微照片）间集中到液相（共熔相)。注意在凝固刚开始时，溶质浓度升高，同时剩余溶液冰点降低。标尺（Bar）=26.7 μm。

金属催化剂的异质混合物存在下进行的，反应物初始浓度低至 10^{-6} mol/L。通常是利用 Mg（Ⅱ）/Pb（Ⅱ）混合物，但 Pb（Ⅱ）离子自己在与单体的比值仅为 1∶20 时也能催化多聚化反应[19]。低浓度的活性寡聚物可有效结合，产生较长多聚物，或通过添加新的活性单体来进行延长。有趣的是，只有寡聚物正在形成二级结构时，单体的添加才能够生产出长于 30～35 聚物（Monnard and Szostak，未发表)，这个事实可能意味着反应过程偏向能形成一定结构的产物，这会有利于产生更多的 RNA 催化物种。

与在均匀水溶液中相比，在冰水共熔相中产物的区域选择性（3′—5′连接）也更强[18]，且冰水共熔相有利于降低水解速率。在典型反应中，仅有 6%～10%初始 ImpU 浓度会水解为失活的尿核苷单磷酸。相反，在均匀水溶液及蒙脱石上，水解速率与单体浓度成正比，并会被金属离子加速至 600 倍[19]。在单体供应可能不足的环境中，共熔相中的低水解速率对产生和选择较长的寡聚物都是必要的。

10.4.4　脂双层结构

在无水和有水条件的循环变动中，如果热能可以为脱水过程提供活化能，便会驱动磷酸二酯键的合成，而不需要单体核苷酸的化学活化，这个可能刺激了对脂双层结构这种特异微环境的探讨。用两亲双分子层组成囊泡或脂质体，可以作为这样的变动环境的模型。在适宜条件下，两亲分子悬浮在水介质中形成双层结构，即使囊泡脱水，这种分子排布也能得到保留。如果 RNA 单体这样的小分子在脱水过程中存在于水介质内，它们就会被包裹到两亲双分子层内（图 10.8)，其排布结构可能有利于多聚化作用。

这一多聚途径与之前提到的两种有以下不同[20]：①单体没有被化学激活；

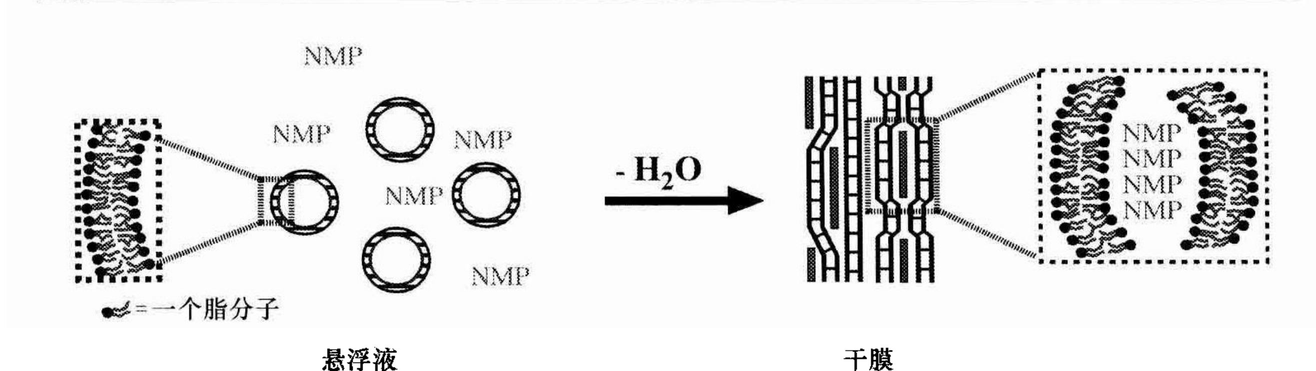


图10.8 脂双层格的形成。RNA单体（NMP，核苷酸5′-单磷酸）与脂质体悬浮液（脂双层为环状和类似梯子的柱状结构）混合。随着混合物的脱水，脂质体融合成多层结构，形成仍具一定湿度的脂双层格，将单体分隔为延伸的分子堆（灰线）。

②没有金属离子做催化剂，因此，RNA连接的形成必然需要额外的能源，如热能；③两亲双层的结构使系统保持在稳定状态；④进程结束时，RNA产物片段将会被分隔开来。我们之后将回到这一点。

在Deamer及其同事的实验中[20]，将溶解在乙醇中的磷脂酰胆碱脂质（这种分子在前生物地球上可能并不存在）注射到含有核苷单磷酸的水溶液中，它们能够自发形成膜结构，然后加热这种悬浮液，在CO_2气流中干燥，形成脂双层膜，在70～90℃间孵育2h，然后将样品再次水化，并使干燥—加热—再水化过程重复7次。

如果样品只含有一种核苷酸单体，腺嘌呤5′-单磷酸（AMP）或尿嘌呤5′-单磷酸（UMP），就能检测到长度为50～100个单体的RNA分子，最多占生成的多聚物重量的10%。最大产量取决于几个实验参数：温度、循环数、悬浮液脱水过程中使用的气体、脂类组分及脂质与核苷酸的摩尔比。UMP和AMP的混合物产生长度较短的多聚物，这与蒙脱石或冰水共熔相中观察到的一致。若没有脂质，而循环数一致，就不能检测到长于四聚体的产物。最后，RNA产物的区域选择性仍然未知，研究者们只能说“产物不能被RNA酶消化”[20]。

总之，这三种RNA单体多聚化的不同途径得到的结果清楚地显示，可以有多种途径去产生能够具有RNA活性的多聚分子，这些多聚物仍包含多种连接方式，这可能对它们的活性和复制产生了很强的抑制作用。

10.5 RNA世界中的信息传递

活性RNA分子的复制（图10.9）对于RNA世界的发展是必需的，这是因为它能使RNA数量选择性增加。并且，最初似乎相对不那么精确的复制过程允许产生少量突变，它与选择性一起，可能共同促进进化出新的和（或）更好的催化反应。但是，复制必须足够精确，以保证RNA的功能，这样的过程，也称为模板介导的多聚化，只有在存在一个RNA分子作为模板、单体或短的寡聚物能

够通过碱基配对形成连接并装配其上的条件下才能发生。注意，催化型的 RNA 分子复制时至少会失去部分结构。

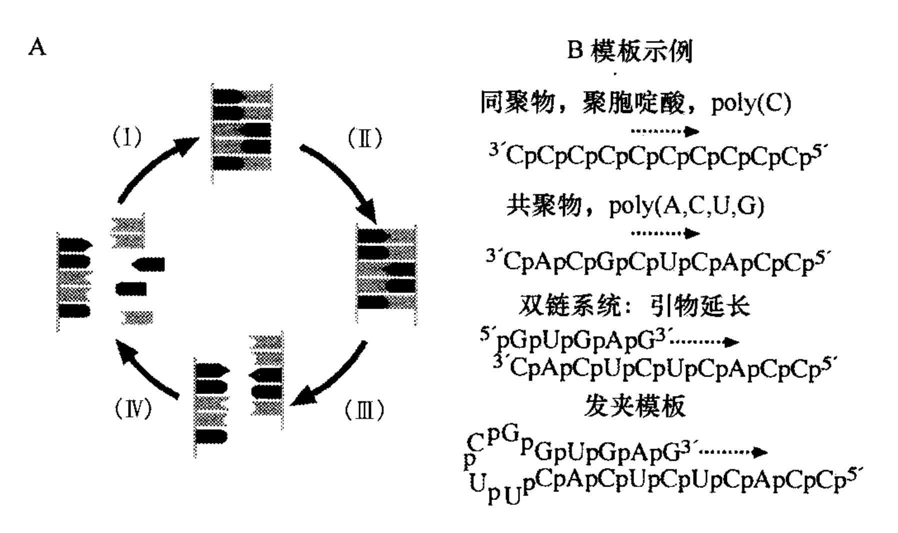


图 10.9　模板介导的多聚化作用的图示。A. 模板、单体和金属离子催化剂混合，并在双层结构（I 阶段）中通过碱基配对进行装配，使单体靠近；通过金属离子催化剂合成第二条链的骨架，生成一条完整的互补链（II 阶段）；双链产物分开（III 阶段），成为其自身的模板，其互补链生成最初的那个序列（IV 阶段）。B. 模板示例。箭头显示多聚化的方向。

在模板介导的多聚化过程中，有两个方面影响单体和寡聚物的连接能力：①杂交系统的稳定性，随互补链长度增加和（或）更长寡聚物的利用而增强；②链上相邻单体或寡聚物的相似性可以使它们自我排列，促进多聚化。

检测单体自发浓缩的环境也用于检测模板介导的 RNA 复制过程，值得注意的是脂双层结构是个例外，目前还未被研究。除了常规化学物质，样品还包含 RNA 多聚物作为模板（图 10.9）。多数反应都在离子强度很高的低温（0℃左右或以下）介质中进行，以分别减少模板和单体的排斥作用（都带负电），促进反应物间的碱基配对。

10.5.1　均匀水溶液中模板介导的多聚化作用

在这种环境中，嘌呤在其互补的嘧啶链上能够很容易进行一种有效的非酶、模板介导的多聚化作用。由于核酸模板上的碱基配对不能克服 U 单体的弱堆积，所以两个连续的 A 残基就能完全抑制反应的进行（这意味着需要向互补链中加入两个 U）。并且，实验观察到，模板必须至少含有 60%的 C 残基，以进行有效“复制”[21]，这个结果可能与活性 G 残基最易堆积有关，令产生的系统排列更好，从而促进多聚过程。即使在模板含有大于 60%以上的 C 残基时，互补链能够形成，但仍不能指数增长，富含 C 残基的模板本身可能也不能复制，因为它自己的模板会含大于 60%的 G 残基，而 C 残基将小于等于 40%。利用活性寡聚物可

能部分解决这个问题，因为寡聚物杂交更为稳定，增强多聚化作用。已经被证实，使用两种 RNA 寡聚物、三磷酸盐[22]或活性嘧啶五聚物 [Im (pU)$_5$][23]及更长的 U 寡聚物，分别主要产生 3′—5′连接产物或 2′—5′连接产物（这两篇观察到不同区域选择性的文献使用了不同的反应步骤）。

10.5.2 矿物表面上模板介导的多聚化作用

模板介导的 RNA 多聚化 [poly (U) 上的 ImpA 和 poly (C) 上的 ImpG] 可以被几种矿物质引发，如羟磷灰石（hydroxylapatite）和绿坡缕石（attapulgite），它们是一种黏土[24]。有理由期待，在蒙脱石上，很快就可能观察到新合成的嘧啶与 1-甲基腺嘌呤活性基团（图 10.4）进行模板介导的多聚化反应。它们自身更强的吸附性能够延长黏土上嘧啶模板或单体复合物的半衰期，使模板介导的多聚化反应能够更有效地进行。

10.5.3 冰水共熔相中模板介导的多聚化作用

冰水共熔相似乎有利于核糖核苷酸进行模板介导的多聚化作用。Stribling 和 Miller[25]证明，模板介导的 ImpG 寡聚化过程能够在冰水共熔相中进行。Trinks 及其同事[26]证明，以 poly (U) 作为模板，同时含有 2-MeImpA 的溶液在一年内产生的 RNA 长度可大于 100 个单元。−25℃时，在 poly (A) 上也能生成短的尿嘧啶寡聚物（3～5 聚物）[23]。实际上，我们观察到，在低离子强度的介质中（Monnard and Szostak，未发表），所有 4 种核苷碱基在非竞争条件（即任何时间只存有一种活性核苷碱基）下，在 RNA 发夹模板上的初始延长速率（图 10.9）显然取决于单体与模板的碱基配对，它们与非 Watson-Crick 系统类似（在一个数量级内），或最多高 1.5 倍。比起在溶液中，该速度提高率可能不大，但需要注意的是，在均匀水溶液中不能发生自发浓缩或寡聚物延长，而在冰水共熔相的自发浓缩反应中却总能观测到有效的寡聚化作用。有趣的是，单体自发浓缩可被介质中溶解的盐强烈抑制[27]，而模板介导的多聚化作用在低和高离子强度介质中都能进行。

10.5.4 模板的连接异质性

具有不同连接方式的 RNA 寡聚物可作为其自身复制的模板。可以预期，这种模板的混合连接是其进行有效 RNA 复制时不可逾越的障碍。但是，这种预期只是部分正确，因为在蒙脱石上[28]，以 2′—5′连接的寡（C）为模板，2-MeImpG 成功进行了模板介导的多聚化作用，同时产生了 3′—5′和 2′—5′连接的产物。最近，Sawai 及其同事研究结果表明，纯的 2′—5′连接的 RNA 四聚物仍能在纯的 3′—5′连接的十聚物模板上装配，生成八聚体产物，而纯的 3′—5′

四聚物在纯的 2′—5′十聚物模板上也如此，不过其装配效率不如单一连接（模板和寡聚物）的系统[29]。

总之，对模板介导的 RNA 多聚化的研究结果，为由活性单体或短的寡聚物生成生物多聚物的有效复制的条件提供了新鲜视角。注意，利用寡聚物时，复制的精确性可能不如活性单体。研究结果同时也显示，即使效率比单一连接低，但具有混合连接的 RNA 也能够作为模板，只是目前尚没有实现扩增反应。

10.6　RNA 单独作为向细胞生命进化的原动力

人们仍然可以质疑，要产生一个 RNA 世界，RNA 本身是否足够？假设存在一对 RNA 聚合酶（一个模板和一个复制子，不过在这里也包括连接酶），它们自身是否足以产生 RNA 生命？事实上，单独考虑 RNA 世界的出现和进化似乎是很离谱的[30]。我们所知的生命提示，从很早开始就有了独立的代谢结构，这对创造基于 RNA 的前细胞是必需的，仅仅由于两者的同时出现和相互反馈调节才使得它们存活了下来。例如，独立结构（使得分子与周围环境隔离开来的结构）对保证 RNA 多聚化的过程（如在脂双层结构中描述的那种多聚化过程，在最后的再水化后，会自发地再次装配成囊泡单元，部分包裹 RNA 产物）、高浓度反应物的积累和保持及特定 RNA 的单独进化肯定也是必需的。

10.7　结论及前景

通过非酶多聚化进行的真正 RNA 扩增，是 RNA 世界产生的奠基石，但其过程仍很神秘。不过，对于理解 RNA 片段如何从一群活性单体中出现及随后的复制，已经有了一定的进展。研究表明，有几种不同途径可以支持这类反应，生成 30～50 个单元的分子，其长度符合催化活性所需。而且，RNA 链的复制既可通过单体，也可通过寡聚物进行。

这些系统仍然有弱点，其中最重要的是初始化合物的纯度，因此，接下来的工作应当是做出模型，而不仅仅停留在字面表述，例如，在黏土上发生的反应对单体活化基团的性质极度敏感；冰水共熔相中的反应效率似乎取决于溶剂的总浓度，包括那些不参与多聚化过程的；在脂双层途径中，反应混合物中没有金属离子在场，若有的话他们能够加速 RNA 在高温下的分解，而有机溶剂也可能破坏双分子层的自我装配。

杂质或非最佳条件将导致较慢的多聚化过程。如果 RNA 产物不会被迅速分解，那么这还是可以接受的。因此，功能性 RNA 全部多聚化的时间应当比一个完整多聚物水解的时间要短。虽然如此，研究者们已经揭示了三条 RNA 多聚化的途径，并可能有未发现的方案。不同的途径可能同时促进 RNA 多聚化，足以利用更多样的环境条件。

我们仍未证实原始地球上由有机分子从头发展出一个 RNA 世界，仍有一些不可避免的问题亟待解决。①仍需发现一种 β-D-核糖核苷酸的前生物合成方式，不论是不是活化核苷酸，这个过程不仅应该高效，而且应具有足够选择性，由包含庞大数量有机分子并可能妨碍合成的混合物中得到大量纯的单体。②在前生物合成的 RNA 池中是否产生具有催化活性的 RNA 仍未被证明，即使不考虑这些 RNA 是否具有复制能力。③仍需阐明其他前生物前体，如独立单元空间，在发展成熟 RNA 生命过程中的作用，因为它们可能有辅助的作用。

考虑这些问题时，不能忘记 RNA 可能并不是最早的前生物信息和催化系统。可能曾经存在一种基于自我装配网络的系统，具有相互联系的催化反应，而没有 RNA 或其替代物介入，它可能起始了一系列的进化步骤，带来生命的产生。尽管如此，这种代谢世界的进化最终一定会产生一个遗传信息或催化-多聚物系统，与基于 RNA 世界类似，RNA 是现今 DNA-RNA-蛋白质世界的前体。这种情况的一个可能优势，是为 RNA 单体的合成提供有效的催化途径。

因此，不管人们是否相信 RNA 世界的存在，或进一步考虑 RNA 是否真的是第一种信息或催化载体，在没有复杂代谢反应的情况下，特别是目前没有其他任何有实验依据的选择时，非酶催化的 RNA 多聚化和复制的研究，可以为探索多聚信息或催化系统产生的必要条件提供难遇的机会。

致谢

在此，我由衷地感谢 Dr. C. L. Apel，Dr. M. S. Declue，Dr. H. J. Cleaves，Dr. H. -J. Ziock 和 H. Fellermann 以及他们的宝贵建议。此项研究得到了来自 Los Alamos National Laboratory 的 Laboratory-Directed Research and Development Program 项目的支持，尤其是 Protocell Assembly Project 的支持。

参考文献

扩展阅读

1. Bloomfield VA, Crothers DM, Tinoco Jr I. Nucleic acids: Structures, properties and functions. Sausalito, California: University Science Books 2000.
   • 该书描述了核酸性能。
2. Orgel LE. Prebiotic chemistry and the origin of the RNA World. Crit Rev Biochem Mol Biol 2004; 39:99-123.
   • 最新有关RNA世界出现的前生命化学。
3. Cech TR, Atkins JF. RNA-World. 3rd ed. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 2005.
   • 该书覆盖了RNA世界有关RNA功能的多方面的研究。
4. Müller UF. Re-creating an RNA world. Cell Mol Life Sci 2006; 63:1278-1293.
   • 综述基本RNA功能的出现及RNA “生物”。
5. Shapiro R. Small molecule interactions were central to the origin of life. Q Rev Biol 2006; 81:105-125.
   • 该综述挑战RNA在古地球上的自发出现，并描述一个纯催化体系的“代谢世界”模型。

专题文献

6. Crick FHC. The origin of the genetic code. J Mol Biol 1968; 38:367-379.
7. Orgel LE. Evolution of the genetic apparatus. J Mol Biol 1968; 38:381-393.
8. Gilbert W. The RNA-world. Nature 1986; 319:618.
9. Steitz TA, Moore PB. RNA, the first macromolecular catalyst: the ribosome is a ribozyme. Trends Biochem Sci 2003; 28:411-418.
10. Zubay G, Mui T. Prebiotic synthesis of nucleotides. Origin Life Evol Biosphere 2001; 31:87-102.
11. Eschenmoser A. Chemical etiology of nucleic acid structure. Science 1999; 284:2118-2124.
12. Böhler C, Nielsen PE, Orgel LE. Template switching between PNA and RNA oligonucleotide. Nature 1995; 376:578-581.
13. Horthota AT, Szostak JW, McLaughlin LW. Glycerol nucleoside triphosphates: Synthesis and polymerase substrate activity. Org Lett 2006; 8:5345-5347.
14. Kanavarioti A. Preference for internucleotide linkages as a function of the number of constituents in a mixture. J Mol Evol 1998; 46:622-632.
15. Ertem G. Montmorillonite, oligonucleotides, RNA and the origin of life. Orig Life Evol Biosph 2004; 34:549-570.
16. Ferris JP. Montmorillonite-catalysed formation of RNA oligomers: the possible role of catalysis in the origins of life. Phil Trans R Soc B 2006; 361:1777-1786.

17. Miyakawa S, Ferris JP. Sequence- and regioselectivity in the montmorillonite-catalyzed synthesis of RNA. J Am Chem Soc 2003; 125:8202-8208.
18. Monnard PA, Kanavarioti A, Deamer DW. Eutectic phase polymerization of activated ribonucleotide mixtures yields quasi-equimolar incorporation of purine and pyrimidine nucleobases. J Am Chem Soc 2003; 125:13734-13740.
19. Kanavarioti A, Monnard PA, Deamer DW. Eutectic phases in ice facilitate non-enzymatic nucleic acid synthesis. Astrobiology 2001; 1:271-281.
20. Rajamani S, Vlassov A, Coombs A et al. Lipid-assisted synthesis of RNA-like polymers from mononucleotides. Origin Life Evol Biosphere 2008; 28:57-74.
21. Joyce GF. Non-enzymatic template-directed synthesis of informational macromolecules. Paper presented at: Symposia on Quantitative Biology, Cold Spring Harbor 1987; 52:41-51.
22. Rohatgi R, Bartel DP, Szostak JW. Nonezymatic, template-directed ligation of oligoribonucleotides is highly regioselective for the formation of 3'-5' phosphodiester bonds. J Am Chem Soc 1996; 118:3340-3344.
23. Sawai H, Wada M. Nonenzymatic template-directed condensation of short-chained oligouridylates on poly(A) template. Origin Life Evol Biosphere 2000; 30:503-511.
24. Schwartz AW, Orgel LE. Template-directed polynucleotide synthesis on mineral surfaces. J Mol Evol 1985; 21:299-300.
25. Stribling R, Miller SL. Template-directed synthesis of oligonucleotides under eutectic conditions. J Mol Evol 1991; 32:289-295.
26. Trinks H, Schroder W, Biebricher CK. Ice and the origin of life. Orig Life Evol Biosph 2005; 35:429-445.
27. Monnard PA, Apel CL, Kanavarioti A et al. Influence of ionic solutes on self-assembly and polymerization processes related to early forms of life: Implications for a prebiotic aqueous medium. Astrobiology 2002; 2:139-152.
28. Ertem G, Ferris JP. Template-directed synthesis using the heterogeneous templates produced by montmorillonite catalysis. A possible bridge between the prebiotic and RNA-worlds. J Am Chem Soc 1997; 119:7197-7201.
29. Sawai H, Wada M, Kouda T et al. Nonenzymatic ligation of short-chained 2'-5'- or 3'-5'-linked oligoribonucleotides on 2'-5'- or 3'-5'-linked complementary templates. Chem Bio Chem 2006; 7:605-611.
30. Szostak JW, Bartel DP, Luisi PL. Synthesizing life. Nature 2001; 409:387-390.

11　核酶与代谢的进化

Randall A. Hughes，Andrew D. Ellington

11.1　引言

在前生物世界中，不存在已知生命体，但是在这个生化进化的过程中，产生了携带信息和行使催化功能的分子。对现代生命体的比较分析显示，多数生物信息都以核酸的形式储存（DNA 和 RNA），而蛋白质负责多数催化反应，但为什么总是这样呢？显然，促进合成反应的催化剂及遗传物质的解码都是必需的，而遗传物质必须编码这些催化剂的合成，这是先有鸡还是先有蛋的问题最根本的来源。

我们难以确切知道最初可能有哪些分子。催化作用的出现，以自身作为反应物或催化其他反应，将可利用的前体转化成产物，丰富了原始世界的化学多样性，最终形成自动催化循环。但是，催化作用在一定程度上还受天然材料的可利用性及催化剂可能发生的降解限制。在这种情况下，维持化学多样性是极端困难的，因为循环建立和新催化剂的获得都要受到相当随机的环境变动的影响。这个问题的一种解决方式是发展出一套系统来储存用催化剂构建的编码信息。已知在现代生物学中，DNA 和 RNA 行使这种功能，系统发生学数据显示，这些化合物相当古老，而核酸的结构单元有可能存在于原生汤中，核酸多聚物携带的信息及随后传递到有用的催化剂的过程，使得早期基因组及携带这些基因组的生命体具有巨大的进化优势。更为重要的是，遗传系统能够突变，经历选择，从而适应不同环境。

关于前生物世界的构成，早期的假说认为，简单的催化循环包含携带信息的 RNA 及能够催化的蛋白质。但是，20 世纪 80 年代初期，Thomas Cech 发现，四膜虫前 rRNA 中的 Ⅰ 类内含子能够催化其自身从 rRNA 转录子上的切除过程[4]，而 Sydney Altman 发现，核酸酶 P 的 RNA 部分能够催化 tRNA 转录子前体的裂解[5]。这些发现推进了核酸可能在前生物世界中扮演双重角色这个概念，并有可能解决经典的鸡和蛋谁先出现的问题。一得到充实，这个想法最终就成为“RNA 世界”假说[6]。这个观点认为，前生物世界可能由基于核酸的简单复制子构成。早期 RNA 世界中，一旦形成这个基因型与表现型之间的联系，其他 RNA 催化剂（核酶）就能够演化，催化一些反应，包括早期生物合成途径。最终这些核酶精细到一定程度，发展出翻译装置，然后核酶就会逐步被其相应蛋白

质取代。

尽管现今只有少量天然核酶存在，但大量代谢反应的合成都能追溯到 RNA 世界中产生的核酶。多数的现代辅因子，如 ATP、NAD，甚至叶酸，在某些方面都是基于核糖核苷酸的（图 11.1）。White 及一些人认为，一旦有一定量的酶/核酶开始利用特定的辅因子，只要生命体的适应性没有极度下降，这种辅因子就不会改变（所谓的“多用户原则”）[7]。如果对于每一套基于特定辅因子发展出的催化剂，其辅因子都不变，那么我们可以认为，它们是细胞中所有生物结构及古老代谢环节中进化最为缓慢的。将这个分析推广，可以假设，这些辅因子最初是被 RNA 催化剂引入和利用的，并被转移进入现代，这是因为进化出新的催化剂比为催化剂重新进化出新的反应要容易一些[8]。这对于现代的蛋白质酶来说显然也是正确的：蛋白质序列、结构和底物都很容易改变，但蛋白质催化的反应和利用的辅酶相对固定。例如，多用户原则很容易解释为什么脱氢酶通常通过它们的腺苷酸结合 Rossman 折叠相联系，而不是它们的底物结合域。

因此，问题就变成早期核酶催化哪些反应，以及这些反应如何促进生命的起源和进化。在本章中，我们将讨论一些直接由进化实验得到的可被核酶催化的反应，以及它们如何能够帮助我们理解前生物化学和 RNA 世界。

11.2 复制反应中的核酶

在假想的 RNA 世界中，最早的基因组应该是简单的核酶，能够利用所得材料进行自我复制。相应的，最早且最简单的复制机制应该是通过 Watson-Crick 碱基配对，以自身为模板进行连接。因此，短的寡聚核苷酸的模板复制实际上是从三聚物或更短的底物开始的[9,10]。假设这样的机制能够使短寡聚核苷酸前体初步聚集，而它们经过扩展和复杂化反应能生成一些催化活性更高的较长寡聚物。尤其是，复杂的 RNA 结构将比简单模板连接效率更高。

要支持这些假说，我们必须超出已知的天然核酶的范围，因为它们数量很少，而且已经进化了很多，主要催化 RNA 加工反应，而不是复制反应。定向进化可以创建早期催化剂的副本，为再现 RNA 世界提供绝佳方法。核酶经过进化，能够在连接起始点利用 5′-三磷酸 RNA 作为底物，催化形成 3′—5′磷酸二酯键，从而有力证明，利用定向进化手段可以研究起源问题[11]，这个反应与蛋白质连接酶和聚合酶催化的反应十分类似，不过核酶催化比其相应的蛋白质酶要慢得多。该相似性已被进一步扩展，选择出了利用核苷三磷酸盐作为底物，具有聚合酶活性的核酶[12]，孵育 4d 以后，这种核酶能够将 RNA 引物退火延伸 6 个核苷酸，达到 92%的准确度。但是，要完成第二条链，这种核酶最后必须自我展开，因此大大限制了这种核酶催化自我复制的能力。为了解决这个问题，Bartel 及其同事试图构建出一种能够识别分离（反式）的模板：引物复合体的聚合酶[13,14]。

图 11.1 核苷酸类辅因子。许多辅因子都是由腺苷酸衍生而来的。R 是多种辅因子与 ADP 骨架的结合位点。此外，许多核糖辅因子的 3′-羟基（Y）或 2′-羟基（X）都能被磷酸化。

通过复杂的选择机制，他们能够选出一些可以促进催化的反式原始核酶，这些核酶可在 24h 的孵育时间内将 RNA 引物延伸到 14 个核苷酸，且准确度总能超过 98%。不过，由于较低的聚合效率，它们不可能进行更长的合成。

虽然 Bartel 的连接酶和聚合酶可作为一个包含 RNA 聚合酶的 RNA 世界后期的具体表现，但是这样复杂的结构在 RNA 世界早期似乎不会出现。在最早最简单的自我复制寡核苷酸与后期的聚合酶之间，一定存在分子“缺失的一环”（Levy 和 Ellington 假想的“前进的征途”，2001[15]；图 11.2）。这样的缺失环节

有可能是利用寡核苷酸，而不是单体核苷酸作为底物的核酶聚合酶。通过定向进化产生了一些小的核酶连接酶，最近更得出其中一种核酶连接酶的晶体结构，使

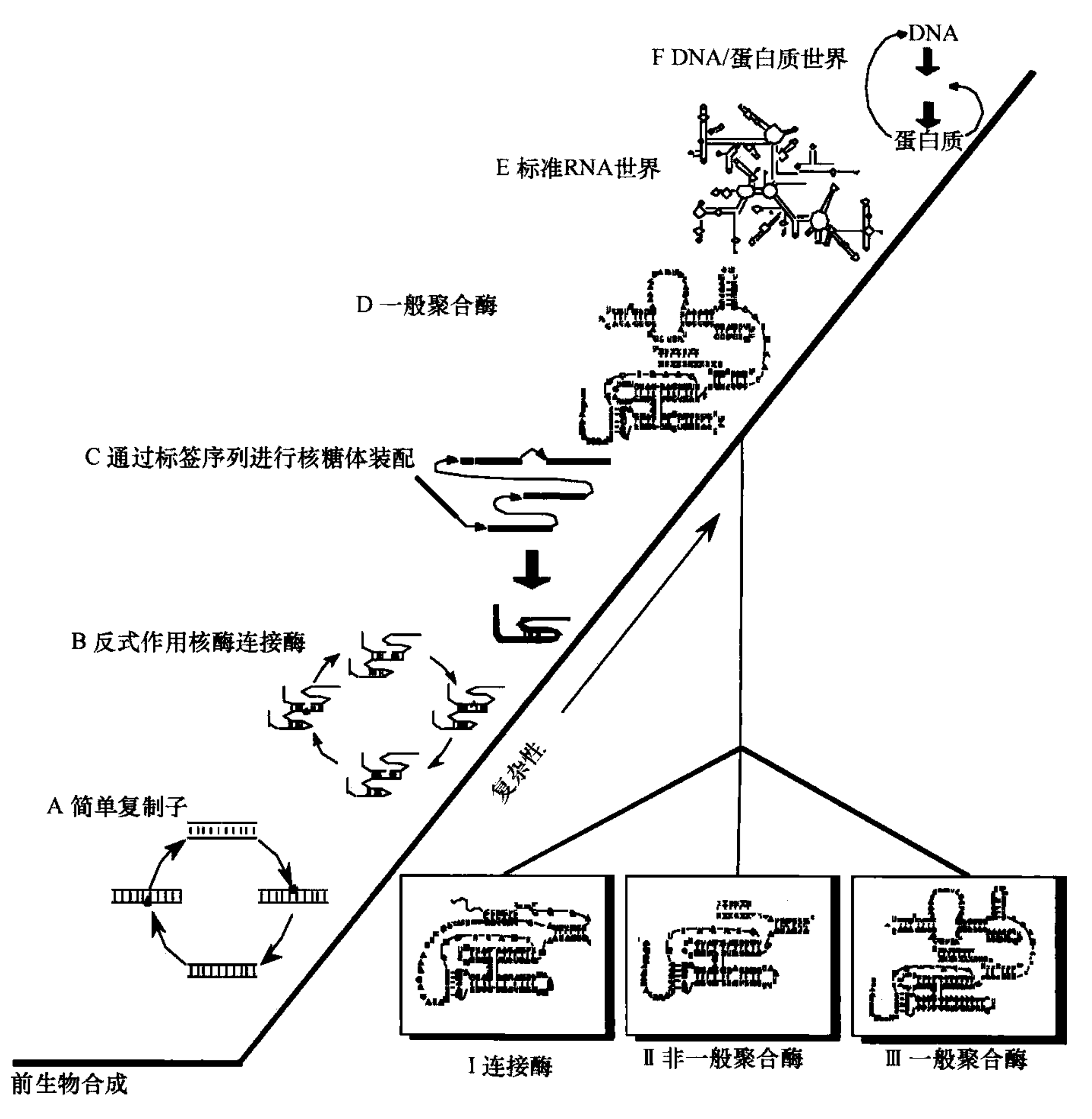


图 11.2　核酶催化过程中的“前进的征途”，这是一种有关原始核酶复制子如何随时间演化，增加复杂程度的观点。A. 简单复制子可能来源于模板介导的寡核苷酸连接；增加的复杂程度会产生反式作用的核酶（B），它能催化自身模板的延长，也能催化其他 RNA 分子。C. 为了避免寄生现象，核酶可能会利用同源序列的较短区域（“标签”）来产生基因组。D. 长链 RNA 的伸展会为模板介导的核酶聚合酶进化提供原料，它能进行基本的复制反应。通过与 Bartel 聚合酶定向进化类似的途径，能产生更好的聚合酶[10]。这个聚合酶由连接酶起始（I），它随后进化成一种非通用的聚合酶，带有模板（II），最终成为更通用的聚合酶（III）。具有更强精确性和持续合成能力的复杂核酶的出现使核酶和基因组能具有更高的序列复杂性，它们反过来又能催化 RNA 世界中可能具有的许多生化反应（E）。最终，翻译途径的出现将建立现代 DNA/蛋白质世界的基础（F）（见彩插）。

我们得以对前生物复制化学略有了解[16]。计算模拟显示，具有更高效率和稳定性的复制子进化存在自举作用（boot-strapped），使早期基因组的大小随时间增加[17]，这进一步支持了“前进的征途”这个模式，这个模式有一个有趣的预言，复制子复杂性的发展与复制子本身的空间距离直接相关。在计算框架内，能够自由扩散的复制子组成并不会带来更高的复杂性，而那些被空间距离分开的却可以。因此，复制子进化可能与基因组的早期细胞化过程联系紧密。

11.3 核酶与辅因子

利用蛋白质做催化剂，一个已知的好处是氨基酸侧链的可用化学多样性远大于 4 种经典核糖核苷酸。氨基酸可以使近中性 pH（组氨酸）时的质子交换、硫醇化学反应（半胱氨酸）和亲核攻击（丝氨酸）反应得以发生。氨基酸也具有多样的理化性质，包含带正电的部分（组氨酸、赖氨酸、精氨酸）和疏水基团（异亮氨酸、亮氨酸、缬氨酸等）。天然核酸极其缺乏这些性质。尽管如此，即使看起来具有这么大的优势，许多蛋白质仍需要辅因子。有趣的是，许多这些辅因子都是基于核苷酸的（图 11.1）。如前所述，这种关联很大程度上意味着这些辅因子是在 RNA 世界中出现并传承的。

核酸辅因子可能早已存在于 RNA 世界中，这是因为泛酰巯基乙胺（CoA 前体）和烟碱（NAD 前体）合成所需的材料都能从前生物途径获得[18,19]。关于辅因子如何被早期核酶利用有不同的模型，但可以想象，作为非特异性连接反应的结果，核酸辅因子应该能够与核酸催化剂共价连接。实际上，已经知道Ⅰ类核酸内含子能够使自身与辅因子（NAD^+）及辅因子类似物（脱磷 CoA）共价连接，它与剪切初始阶段向 RNA 中插入鸟苷的一般机制一样[20]。在氨基酸和翻译反应的建立阶段，类似的机制可能也起了作用，我们将在以下部分讨论。

作为共价连接的辅因子如何增强核酶功能的例子，定向进化实验能够将利用 NAD^+，具有氧化还原活性的醇氧化核酶从一群随机序列中选择出来[21]。同一种核酶也能催化逆反应——将苄基乙醛还原为苄基乙醇[22]。这项研究还有另外两个有意思的结果：第一，选择出的乙醛脱氢核酶对 Zn^{2+} 有绝对依赖，蛋白质酶也有相同依赖；第二，可以通过 NADH 到 FAD 的非催化转化产生的 NAD^+ 来氧化乙醇，意味着在最早的代谢途径中，也可能出现与现代生物化学类似的偶联氧化还原反应。

定向进化也能用于生成那些利用非共价结合辅因子的核酶。向一群随机序列中添加一个 RNA 腺苷结合域（这个 RNA 等价于蛋白质中的 Rossman 折叠），通过在硫醇柱上捕获自激活酶（self-kinases），能够选择出一种核激酶，它能将可溶 ATP-γS 中的硫代磷酸盐催化转化为核酶的 5′-羟基[23]。其中一种选出的核酶能够催化与其 5′端一样的寡核苷酸进行多重反转磷酸化，这个反应与多核苷

酸激酶这种蛋白质酶的催化反应属于同一种。

原生汤中，可利用的少量辅因子前体可能不仅能带来依赖辅因子的核酶进化，也可能带来辅因子生物合成途径的进化。令人满意的是，定向进化实验证实，核酶能够催化由前体形成辅因子的反应。最初由 Yarus 和同事发现的一种焦磷酸盐转移酶（cappase），其底物特异性非常低[24]。实际上，任何包含磷酸盐的分子都能自动结合到核酶的 5′-三磷酸上，取代焦磷酸盐。最初显示，这种核酶能够自动结合辅因子，这与上述Ⅰ型内含子的实验类似。其他核酶将前体结合到 5′腺苷酸上，形成一般辅因子的标准双磷酸键[25]（图 11.3），例如，由 4′-磷酸泛酰巯基乙胺、烟碱单核苷酸（NMN）、黄素单核苷酸（FMN）前体分别合成 CoA、NAD 和 FAD 的核酶。这些实验再次支持了辅因子能够与早期 RNA 催化剂共价结合的假说，这种结合可能在末端。随着代谢作用的扩展，以及对扩散性辅因子需求的增加，共价连接的辅因子可能分离开来，而核酶将发展出利用 ATP 作为反式底物来合成辅因子的反应。辅因子最初到底是核苷酸的附加物，还是携带有具有分子把手的扩散性分子，仍有待商榷，但清楚的是，在早期 RNA 世界中，已经具备合成和利用辅因子的能力[26]。

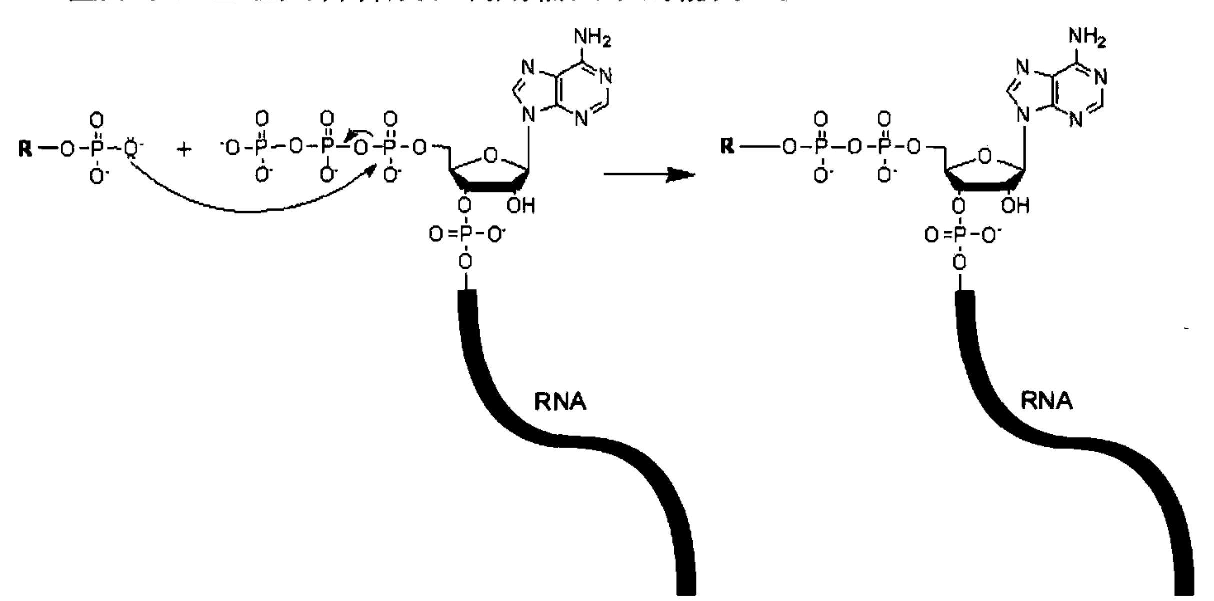


图 11.3　辅因子合成的核酶机制。实际上，任何带有磷酸盐的化学基团（R）都能与“cappase”核酶上的 ATP 端反应，将其自身与核酶的 5′端相连。R 基可以为图 11.1 中显示的任何辅因子部分。

定向进化实验已被广泛用于探索与原始生物合成相关的核酶反应，其中许多都需要辅因子。对核酶化学及辅因子的更详细回顾，建议读者参看额外阅读文献。

11.4　核酶、氨基酸和翻译的出现

高效翻译途径的出现是 RNA 世界生化进化的高峰，同时也是黎明。高效的

翻译能够产生更长的肽链，进行更有效的催化。由于蛋白质催化剂比起核酶有很多优势，一旦它们能从可遗传的基因组 RNA 中翻译出来，它们就能进化并取代效率较低的相应 RNA。但尽管如此，高度进化的核酶起初仍比初期未进化的蛋白质效率更高。因此，似乎最早的翻译产物是一些较短的多肽，作为必要的辅因子帮助 RNA 行使化学功能（从而产生最早的核糖核蛋白）。

这种观点可能能够解释分子生物学上一个长期以来的谜题：为什么先进化出翻译途径？一个优势是，翻译反应可能是那个著名的悖论，即预期进化的典型例子（即在假设的早期细胞中，核酶是初期催化剂，但不知怎么回事进化出了可以取代自己的途径）。即使最先翻译出来的多肽增强了核酶介导的催化作用，但是首先一定存在直接而明显的优势去增强催化作用，而不是发展出一条复杂途径。一个可能的解释是，因为在原生汤中能够获得氨基酸，核酶刚开始时可以利用这些简单而易得的化合物作为辅因子，催化精确性和复杂性的进一步增强会驱动氨基酸以串联及链的形式结合和多聚化。翻译出的多肽如何与 RNA 催化剂共同行使功能的一个例子是，依赖多肽的核酶连接酶活力可以增强 18 000 倍以上[27]。这些实验也证明，除了能作为辅因子外，多肽也可能作为别构剂来调节 RNA 世界中的代谢反应。

这种假设只回避了 RNA 催化剂如何利用氨基酸这个问题。我们之前关于辅因子的讨论提供了一种可能解释：氨基酸可能与核酶本身共价或非共价结合。两种可能性都有实验支持。Roth 和 Breaker 选择了一种利用非共价结合组氨酸作为辅因子的脱氧核酶，催化一种 RNA 底物的裂解[28]。在这些实验中对组氨酸的选择尤其有趣，因为组氨酸是唯一一种能在生理 pH（约 7.0）进行普通酸或碱催化反应的氨基酸。组氨酸咪唑侧链的这种功能在蛋白质中显然广泛存在，在核酸催化剂中尤其有用，因为它们在中性 pH 时不易进行质子转移。也可能利用其他氨基酸辅因子，因为定向进化实验显示，RNA 适体可以得到演化，与大量氨基酸特异性结合（见 Majerfeld et al.，2005[29]及其参考文献）。

但是，氨基酸- RNA 复合体的结合常数相当弱，这意味着共价结合对催化效率的影响更加强烈。氨基酸能够通过相对简单的化学反应与核苷酸直接结合[30]，这些核苷酸随机插入催化剂，至少会带来部分可增强催化能力的变异，这也能帮助建立选择压力，带来原始氨基酸生物合成途径和翻译过程的发展，以及遗传密码的开端，这是由 Wong 提出的假说[31]。但是，为了确定这种催化增强作用在每代中都能遗传，有必要进化出一些机制，将特定氨基酸（或密切相关的氨基酸）结合到核酶的特定位点。关于这一点，Szathmary[32]提出，可以利用连有核苷酸把手的氨基酸。这个想法与催化剂结合利用核酸辅因子类似。核苷酸把手或配体为酶特异性捕获氨基酸辅因子提供了更方便的机制。实际上，带有寡核苷酸把手的辅因子能够与核酶上的特定位点直接杂交。通过某些途径，连有寡核苷

酸的氨基酸可能装配成现代 tRNA 的反密码子茎环，作为进化过程的一步，带来多肽的编码及早期遗传密码的出现[32]。

假设有充足并有力的进化压力能使翻译出现，但庞大复杂的核酶，如核糖体中有一部分行使核酶功能，是如何从相对简单的反应机制中产生，仍是一个问题（Woese 早在 20 世纪 70 年代提出）。定向进化再一次用实验告诉我们在遥远的过去可能发生了些什么，因为较简单的核糖体能催化蛋白质合成的所有所需反应。

在现代翻译中，氨基酸聚合成蛋白质，首先形成高能中间物，然后作为编码底物，最终形成肽键。高能氨酰基腺苷酸中间物通过一类称为氨酰基-tRNA 合成酶（aaRS）的酶及其同源 tRNA，使用 ATP 进行合成（图 11.4，反应 1）。然后 aaRS 在氨基酸的羧基基团和其同源 tRNA 的 3′-OH 间形成酯，生成酰化 tRNA（图 11.4，反应 2），从而将氨基酸转移到同源 tRNA 上。酰化 tRNA 可通过核糖体的肽基转移活性（图 11.4，反应 3）用于蛋白质合成。由于氨基酸和 tRNA 间的酯连接比之后产生的肽键能量更高，这个过程是放能反应，当两个底物在核糖体中接近时，应当能够自发进行。

令人满意的是，定向进化实验显示，核酶能够催化所有这些步骤。Kumar 和 Yarus 的工作证明，能够选择出将 ATP 和羧酸中催化形成乙酰腺苷酸的核酶[33]（图 11.4，反应 1）。离体选择的核酶也能催化随后的氨酰化反应（图 11.4，反应 2）。Yarus 及其同事还能够选择出那些可以利用苯丙氨酸腺苷酸作为起始物，令其自身接受苯丙氨酸氨酰化的 RNA[34]。不可思议的是，选出的一种变体——RNA 77，能够自我用苯丙氨酸氨酰化，而其速度及特异性更超过了现代蛋白质酶的 tRNA 氨酰化[35]。其他的定向进化实验显示，选出的核酶也能催化将其他活性氨基酸转移到它们的同源 tRNA 上，这个反应与蛋白质 aaRS 催化的第二个反应相同（如 Lee 等的工作[36]）。自我氨酰化 RNA 77 的核心催化活性最终能够表现在最小的核酶中，仅有 29 个核苷酸长，尽管其速率比较长的变体有所下降[37]。这些结果表明，RNA 能够催化翻译的三个步骤中的两个，使用的催化剂比现代复合体更简单，暗示着翻译的分子起源可能是相对简单的 RNA 催化剂，后来随时间进化才变得更加复杂。蛋白质合成的第三步也能被核酶催化（图 11.4，反应 3）。例如，最小的 RNA 77 也能催化形成 RNA-Phe-Phe 二肽（图 11.4，反应 3）。Zhang 和 Cech 也选出了具有肽基转移活性的核酶，能够将各种氨基酸形成多肽[38]。要用实验确定翻译途径，下一步是给出序列特异的解码，这将能相对直接地证明氨酰化氨基酸能够结合到寡核苷酸上，如 Szathmary 提出的那样[32]。即便如此，要充实我们对这个进化阶段的理解，仍需进一步实验，可能利用能够以自我为模板形成多肽的核酶，这与 Lee 等已进行的反式酰化实验类似[36]。

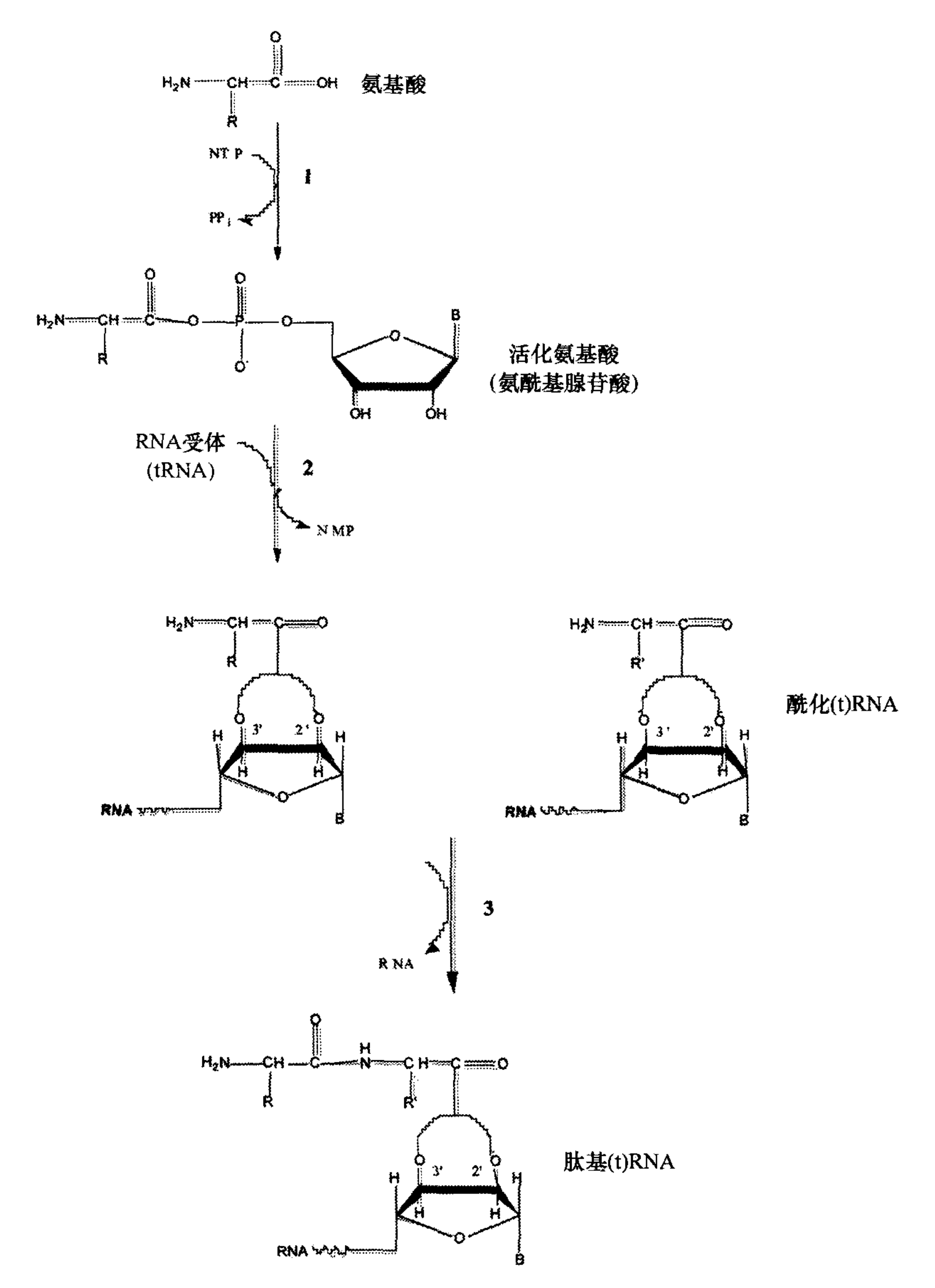


图 11.4 与蛋白质合成有关的核酶催化反应。翻译过程中的许多生化反应也能被核酶催化。①氨基酸活化，形成高能“活性”氨酰基腺苷酸。②RNA 受体上活性氨基酸向羟基的转化。在现代翻译中，受体 RNA 是 tRNA，但是在 RNA 世界中则可能是一种核酶，其功能与现代氨酰基 tRNA 合成酶类似，其附着物可能在 2′或 3′羟基上。③两个酰化氨基酸能反应形成肽键。这个多肽仍与 RNA 连接，正如核糖体上的 P 位点一样。这个步骤的模板还没有通过实验证实，但这种模板可能是现代遗传密码的前身。B 可以是任何核苷碱基。

11.5 结论

自从在 20 世纪 80 年代早期发现了催化型 RNA，定向进化实验已经为早期进化的理论构建提供了实验证据。已经证明，进化的核酶能催化现代代谢中存在的许多反应，在由前生物化学向细胞化学的转变过程中，最终为 RNA 作为模板和催化剂提供了支持。最重要的是，定向进化实验证明，多数现代翻译系统可能由 RNA 世界中较简单的催化剂产生。一旦建立了一个有效的翻译系统，就会为现代生化的进化打开大门。对核酶的进一步研究，将进一步界定哪个过程产生了生命本身，并会帮助我们更好地理解现代生命的最后一个共同祖先，即始祖生命出现和存在的那个生化进化时期。

参考文献

扩展阅读

1. Wilson DS, Szostak JW. In vitro selection of functional nucleic Acids Annu Rev Biochem 1999; 68:611-647.
   • 综述如何进行试管进化取得有功能的核酸。
2. Chen X, Li N, Ellington AD. Ribozyme catalysis of metabolism in the RNA world. Chem Biodiv 2007; 4:633-655.
   • 对核酶催化及RNA世界模型的综述。
3. Joyce GF. The antiquity of RNA-based evolution. Nature 2002; 418:214-221.
   • 出色的有关RNA世界由前生命化学到翻译出现的综述。

专题文献

4. Garriga G, Lambowitz AM, Inoue T et al. Mechanism of recognition of the 5ʹ splice site in self-splicing group I introns. Nature 1986; 322:86-89.
5. Guerrier Takada C, Altman S. Catalytic activity of an RNA molecule prepared by transcription in vitro. Science 1984; 223:285-86.
6. Gilbert W. The RNA world. Nature 1986; 319:618.
7. White HB III. The pyridine nucleotide coenzymes, eds. Everse J, Anderson K, Yu, KS. Academic Press. New York 1982:1-17.
8. White HB III. Coenzymes as fossils of an earlier metabolic state. J Mol Evol 1976; 7:101-104.
9. Sievers D, von Kiedrowski G. Self-replication of complementary nucleotide- based oligomers. Nature 1994; 369:221-24.
10. Zielinski WS, Orgel LE. Autocatalytic synthesis of a tetranucleotide analogue. Nature 1987; 327:346-47.
11. Bartel DP, Szostak JW. Isolation of new ribozymes from a large pool of random sequences. Science 1993; 261:1411-18.
12. Ekland EH, Bartel DP. RNA-catalysed RNA polymerization using nucleoside triphosphates. Nature 1996; 382:373-76.
13. Johnston WK, Unrau PJ, Lawrence MS et al. RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. Science 2001; 292:1319-25.
14. Lawrence MS, Bartel DP. New ligase-derived RNA polymerase ribozymes. RNA 2005; 11:1173-80.
15. Levy M, Ellington AD. The descent of polymerization. Nat Struc Biol 2001; 8:580-82.
16. Robertson MP, Scott WG. The structural basis of ribozyme-catalyzed RNA assembly. Science 2007; 315:1549-53.
17. Szabo S, Scheuring I, Czaran T et al. In silico simulations reveal that replicators with limited dispersal evolve towards higher efficiency and fidelity. Nature 2002; 420:340-43.
18. Dowler MJ, Fuller WD, Orgel LE et al. Prebiotic synthesis of propiolaldehyde and nicotinamide. Science 1970; 169:1320-21.
19. Keefe AD, Newton GL, Miller SL. A possible prebiotic synthesis of pantetheine, a precursor to coenzyme A. Nature 1995; 373:683-5.
20. Breaker RR, Joyce GF. Self-incorporation of coenzymes by ribozymes. J Mol Evol 1995; 40:551-558.
21. Tsukiji S, Pattnaik SB, Suga H. An alcohol dehydrogenase ribozyme. Nat Struct Biol 2003; 10:713-17.
22. Tsukiji S, Pattnaik SB, Suga H. Reduction of an aldehyde by a NADH/Zn^{2+} dependant redox active ribozyme. J Am Chem Soc 2004; 126:5044-45.
23. Lorsch JR, Szostak JW. In vitro evolution of new ribozymes with polynucleotide kinase activity. Nature 1994; 371:31-36.
24. Huang F, Yarus M. Versatile 5ʹ phosphoryl coupling of small and large molecules to an RNA. Proc Natl Acad Sci USA 1997; 94:8965-8969.
25. Huang F, Bugg CW, Yarus M. RNA-catalyzed CoA, NAD and FAD synthesis from phosphopantetheine, NMN and FMN. Biochem 2000; 39:15548-15555.
26. Jadhav VR, Yarus M. Coenzymes as coribozymes. Biochimie 2002; 84:877-88.
27. Robertson MP, Knudsen SM, Ellington AD. In vitro selection of ribozymes dependent on peptides for activity. RNA 2004; 10:114-127.
28. Roth A, Breaker RR. An amino acid as a cofactor for a catalytic polynucleotide. Proc Nat Acad Sci USA 1998; 95:6027-6031.
29. Majerfeld I, Puthenvedu D, Yarus M. RNA affinity for molecular L-histidine: genetic code origins. J Mol Evol 2005; 61:226-235.
30. Robertson MP, Miller SL. Prebiotic synthesis of 5-substituted uracils: A bridge between the RNA world and the DNA-protein world. Science 1995; 268:702-705.
31. Wong JT. Origin of genetically encoded protein synthesis: a model based on selection for RNA peptidation. Ori Life Evol Biol 1991; 21:165-176.
32. Szathmary E. The origins of the genetic code: amino acids as cofactors in the RNA World. Trends in Genetics 1999; 15:223-229.
33. Kumar RK, Yarus M. RNA-catalyzed amino acid activation. Biochemistry a 2001; 40:6998-7004.
34. Illangasekare M, Sanchez G, Nickles T et al. Aminoacyl-RNA synthesis catalyzed by an RNA. Science 1995; 267:643-647.
35. Illangasekare M, Yarus M. Specific, rapid synthesis of phe-RNA by RNA. Biochemistry 1999; 96:5470-75.
36. Lee N, Bessho Y, Wei K et al. Ribozyme-catalyzed tRNA aminoacylation. Nat Struc Biol 2000; 7:28-33.
37. Illangasekare M, Yarus M. A tiny RNA that catalyzes both aminoacyl-RNA and peptidyl-RNA synthesis. RNA 1999; 5:1482-1489.
38. Zhang B, Cech TR. Peptide bond formation by in vitro selected ribozymes. Nature 1997; 390:96-100.

12　前细胞进化：囊泡与原始细胞

Pasquale Stano，Pier Luigi Luisi

12.1　引言

活生命体都是由细胞组成的。多细胞生命由多种特化细胞构成，具有空间组织，并与其他细胞进行复杂的相互作用；同时，单细胞生命只由一个细胞组成，这个细胞必须完成维持生存和繁殖的所有必需功能。根据所有生命的细胞形式，以及在细胞水平之下不存在生命这两点，细胞可以说是生命体最令人惊异的特质。

如果我们看一个单独的细胞，如细菌，我们眼前最明显的事实是：生命是一个与环境很不相同的独立体系。换句话说，细胞具有清晰的界限，将其内部与外部隔开。

观察这两个区域的相互作用，我们也能得到细胞生命的蓝图——通过现象学途径。首先，细胞是活的，但在其任意一个组成中都不能发现“生命”，倒不如说，“生命”更像是一种集体性质，来自于细胞组分和过程的自我组织，这是一种分散式的、突现的性质。从历史上看，一些作者意识到，由非生命到生命的关键转折就是细胞的形成，或者更一般地说，是结构单元的形成。Harold J. Morowitz 在他《细胞生命的起源》（*Beginning of Cellular Life*）一书[1] 中坚定拥护了这一观点，由于不能发现更简单的生命形式，显示出生命结构单元这种普遍性质是不可简化的；并且，他着重强调细胞结构的普遍性，包含将细胞与环境分隔开的质膜结构。

在来到“区室化作用”这个概念之前，首先简要提一下生命起源的现有理论。研究生命起源的科学家们的主要假设是，生命起源于无生命的物质，分子复杂性自发而逐渐地增加，由简单分子到生物单体（如氨基酸、糖类、芳香基、脂类），再到大分子，进而具有活性（所谓的 Oparin-Haldane 假说）。在这个框架内，20 世纪有两类观点主导了科学争论，即先有代谢和先有基因。虽然通常很难将研究者们区分为哪个学派，但像 Alexander I. Oparin，Sidney Fox，Freeman Dyson，Stuart Kauffman，Harold J. Morowitz 和 Günter Wächtershäuser 这样的作者可被宽泛地界定为第一类观点的支持者。其他作者，如 John B. S. Haldane，Leonard Troland，Manfred Eigen，Leslie Orgel，Francis Crick，

Carl Woese，Walter Gilbert 和 Gerard Joyce 则被认为是基因先行假说的追随者。Iris Fry 在她最近关于生命起源的书中，对不同理论选择的来历和进化给出了一个历史性的讨论[6]。

简而言之，这种分裂最终起源于现代细胞中两类生物多聚物的共同出现：蛋白质（酶）和核酸（DNA、RNA），它们循环产生彼此。有多种理论试图解释这两种分子谁先出现，谁是循环的起始。

“结构单元”（compartment）途径可被认为是独立于上述情况之外的，更偏向于区域化反应网络，它不能直接解决代谢或基因的两难问题，但反过来强调了生命起源中独立结构的作用。上述一些作者将他们的观点整合为某种关于结构单元概念的形式，如 Dyson 和 Morowitz。这种途径基本上是一个在自我组织的有界系统中进行的封闭反应（同时也是遗传性的）网络，决定了一系列突现的性质，如选择渗透性、电化学梯度的建立、非均衡过程的维持及最重要的独立细胞的出现。

脂囊泡（脂质体）提供了结构单元的一个基本模型。特定两亲分子分散到水中时，会自发形成脂质体及其他类型的囊泡（图 12.1）。脂分子自我装配形成双层半透膜，封闭形成一个内部水环境，在环境和囊泡内提供了一个疏水屏障。关于生命起源，前生物条件下脂质体单体（即成膜分子）的生成非常重要，这将在以下段落讨论。

结构单元途径植根于 David Deamer，Doron Lancer，Yoiko Nakatani，已故的 Guy Ourisson、Tetsuya Yomo 和我们组的系列研究。最近 30 年，有大量的研究投入到对磷脂囊泡及其他重要的表面活性剂如脂肪酸的探索，并对囊泡性质和反应有了一些重大发现。

在本章中，我们会阐释一些细胞和前细胞进化的基本概念。为了达到这个目的，我们首先讨论怎样定义细胞生命：关于这点，最有力的理论是自生（autopoiesis）学说，接下来会简要介绍。之后我们会以脂质体作为早期细胞的模型，描述一些基本特征，对其中发生的生化反应做一个简要回顾。我们也会审视一下囊泡中一个令人惊异的反应途径，即囊泡的自我增殖，以及在结构单元框架中，可能发生的代谢先行和（或）基因先行的整合。最后，我们会介绍一下“最小的细胞”这个概念及对其的研究（实验室中合成的实际物体）它可能具有最小化的细胞功能。这为理解早期细胞的结构和功能，以及它到现今复杂细胞的进化铺平了道路。

12.2　自生发展和细胞生命的推理

细胞生命是自生发展（autopoieis，希腊语中 auto＝自我，poiesis＝产生）的起始点，在 20 世纪 70 年代由两个智利生物学家 Humberto R. Maturana 和

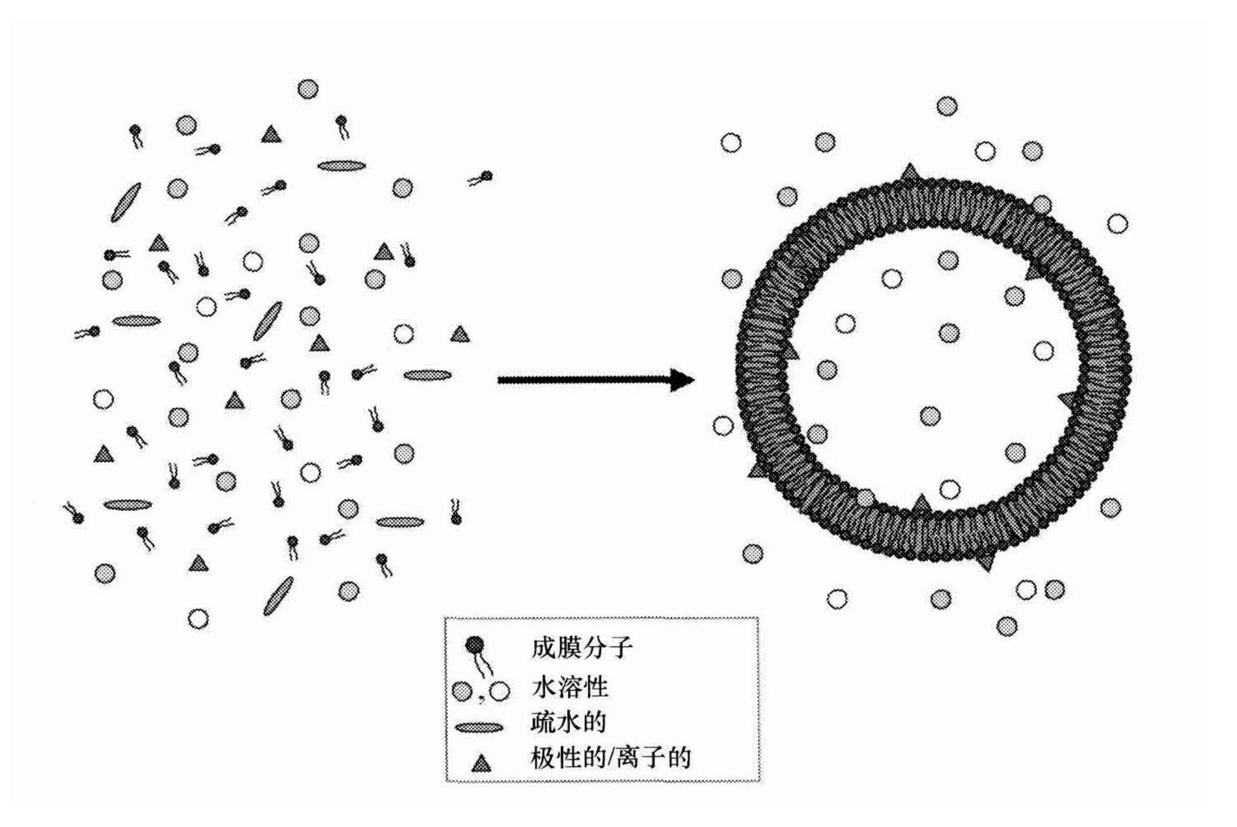


图 12.1　成膜表面活性剂自发自我装配成囊泡，其内腔中的水环境包含水溶性分子。其中任何疏水分子都会定位到膜上；如果在囊泡形成过程中有离子表面活性剂参与，那么带有互补电荷的离子和（或）极性溶质就会吸附到表面。注意：囊泡的形成造成了三相微观混合系统（内部/边界/外部），从而能够建立化学梯度、大分子分泌、选择渗透性和最重要的细胞类独立性的产生。双链表面活性剂及某些单链活性剂能在许多条件下形成囊泡。除了双层膜，所谓的“流星锤”（bola）两亲分子也能形成单层膜。溶质被动内陷到囊泡中的过程不需要额外的能量；表面活性剂自发装配的过程也一样（$\Delta G<0$）。

Francisco J. Varela[2,3,7]提出。自生说并不考虑生命的起源，但回答了“生命是什么?”这个问题，分析了现代的生命体。从这个意义上，自生说是基于现象学的描述性理论，起点是确定了细胞的主要活性是维持其自身的独立性，虽然在其界限内有大量转化反应。这可能要归功于细胞内部再生时，其中所有被转化的组分都得到了重建（包括边界分子!）。因此，简而言之，细胞生命的蓝图植根于一系列能生成所有细胞组分的过程，而这些组分又能反过来生成这些过程。这当然需要来自外部的能量和原料，因此从热动力学上讲，活细胞是开放系统，特点是其自身运作的封闭性，即它包含所有自我组织和复制的信息。

这些实际上都非常简单。自生说认为活细胞在组织和结构两个水平上互补。组织性是所有活细胞不变的性质，它基于其成分的产生，自我组织成反应网络，再生成组分，从而使循环一直进行下去（图 12.2）。然而，细胞结构却可以在不同的细胞或不同的进化阶段间有所变异。

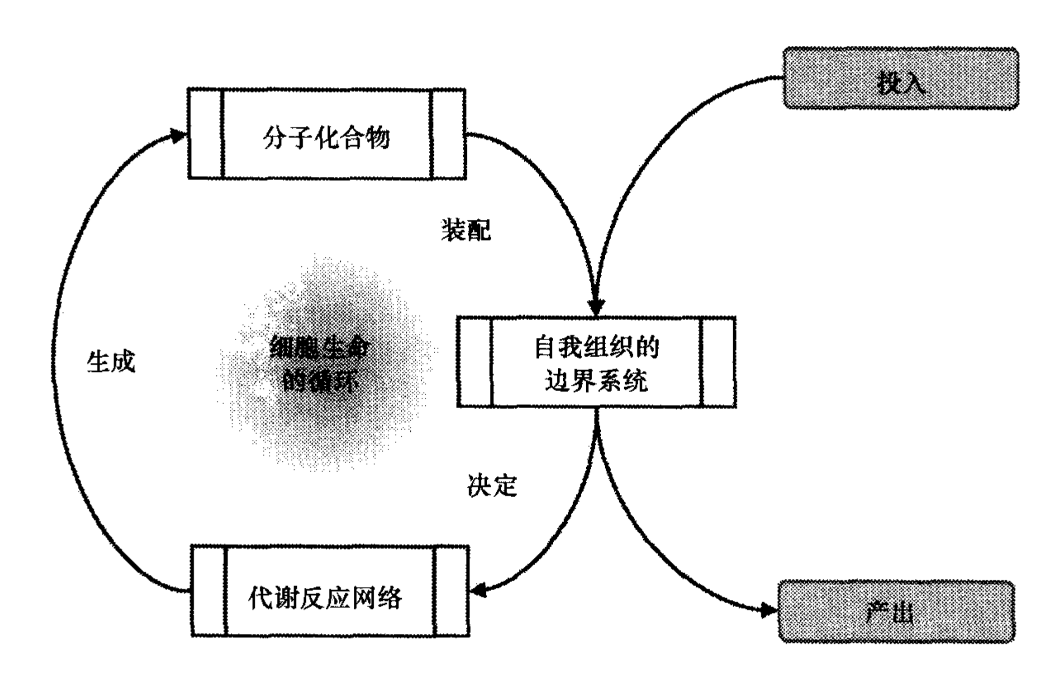


图 12.2　细胞生命的循环。细胞由自生单元开始，是一个自我组织的边界系统，具有特定反应网络，它反过来生成该系统装配所需的分子，又确定了特定反应网络，再如此循环。

与这样的细胞观点相联系，有一些重要概念，如生物自治，社会自生，二级自生和三级自生（取决于生命体的复杂度）等，建议读者浏览 Luisi（2006）[2] 关于自生说的介绍，它也包含历史背景的信息，并讨论了为什么自生说对同时代生物化学没有产生强大冲击。这里也可能提到，为了对生命状态有一个完整概念，自生说也需要认知上的补充，也是由 Maturana 和 Varela[7] 提出的，即与环境间的选择性作用[8,9]。

自生说在细胞水平上与生命的定义紧密联系。实际上，存活这个特性属于细胞整体，而不是单独的某些部分。因此，自生说与引言中介绍的结构单元途径相关，并不是因为它解释了结构单元如何产生，而是因为它强调，对生命整体起源的探索必须关注区域化的系统，而不是单一分子。

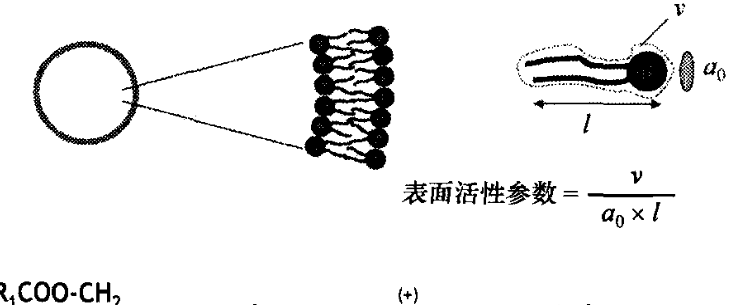
在接下来的段落中，将讨论结构单元的自我增殖，我们会描述自生说是如何为这些研究提供灵感，以及如何在实验室中实现简单的自生化学系统。

12.3　囊泡及其他结构单元

在 Maturana 和 Varela 的工作之前，结构单元是生命基本结构这个概念早已存在。历史上，结构单元的第一个模型是“团聚体”（coacervates）[10]，它是蛋白质和糖类自我装配形成的微滴，能够吸附小分子，将它们转化成原始代谢的形式，这应该是细胞的前体。原始结构单元的第二种模型是由 Fox[11] 建立的，它是

氨基酸多聚化产物形成的“类蛋白质”微球。这两种系统只具有历史意义，因为现今对原始细胞的研究多数都针对囊泡，它与生物细胞更为类似，具有高水平的生化表现。除了囊泡，微胞团（micelle）及反向微胞团也用于构建细胞模型，尤其是考虑到自我增殖的时候（见参考文献 2）。

囊泡为球状结构单元，平均半径范围为 50～1000nm，当适当的表面活性剂分散到水中时形成。它们形成封闭的球状半透膜边界，双层膜，表面活性分子自动装配，使其疏水尾部处于膜内，将亲水头部暴露在水溶液中（图 12.3）。磷脂为最典型的形成囊泡的化合物，而脂类形成的囊泡通常称为脂质体。但是，不是所有磷脂都能形成脂囊泡，例如，头部基团的性质能通过改变“表面活性参数”来影响分子的自我装配性质，它是一种经验值，能用于预测表面活性剂最稳定的装配模式。当表面活性分子呈现出类似圆柱体的形状，$v/(a_0 \cdot l)$ 值约为 1 时，会形成双层膜。



$R_1COO\text{-}CH_2$
|
$R_2COO\text{-}CH_2$
|
$CH_2OP\overset{(-)}{O}_2O\text{-}X$

甘油磷脂

- $X = CH_2CH_2\overset{(+)}{N}(CH_3)_3$　卵磷脂
- $X = CH_2CH_2\overset{(+)}{N}H_3$　磷脂酰乙醇胺
- $X = H$　磷脂酸
- $X = CH_2CH(\overset{(+)}{N}H_3)CO\overset{(-)}{O}$　磷脂酰丝氨酸
- $X = CH_2CH(OH)CH_2OH$　磷脂酰甘油

图 12.3　双层膜的详细情况。两亲分子自我装配成双层结构，亲水极性头部暴露在水环境中。疏水尾端聚集在一起，远离水环境。右上方是“表面活性参数”用基本几何元素（v 是两亲分子的体积，a_0 是“有效”头部表面积，l 是长度）表示的图解。底部为甘油磷脂。通过表面活性参数估计，多数情况下它们形成双层，但有些因素，如盐或酰基链的不饱和性会改变它们的行为方式。两亲分子混合物的性质难以预测。

卵磷脂囊泡与其他磷脂囊泡一样是非常熟悉的，已有生物物理研究来确定它们的形成、稳定性、渗透性、硬度和分子动力学（如分子在双层膜中的扩散、穿膜振动、旋转扩散和翻滚运动）。为了得到囊泡样品的定量概念，可考虑 10 mmol/L 的 POPC（1-棕榈酰-2-油酰基-sn-甘油-3-胆碱磷酸）悬浮液，其中单层囊泡的平均半径为 50 nm。一个单独的囊泡平均由约 90 000 个脂分子构成，自我装配成球形双层膜。1ml 这样的悬浮液包含约 10^{14} 个囊泡，每个囊泡的内部体积

约为 5×10^{-4} fL（1fL=$10^{-3}\mu$L），总的内部体积约为 36 μl（3.6%）。总（内和外）表面积却约为 $1m^2$。因此，亚微米级的脂囊泡具有高的表面积/容积比值，而它们的许多性质都与胶体相同。脂双层的化学性质（从极性上说）使得它能具有选择渗透性。水、甘油、色氨酸、一个普通的蛋白质，以及钠离子在卵磷脂膜中的相对渗透性为 10^9、10^6、10^2、1、1，因此，双层磷脂膜对大分子和带电离子是不通透的，而小的非带电分子有较高通透性。因此，很明显，半透膜建立了对分子化合物选择性摄取或释放的体系，而功能性大分子这样的多聚物，若在囊泡内形成，就能很好保存，并在囊泡内部行使功能。

虽然脂囊泡能很好地构建细胞模型，但可能与生命起源的研究没什么关联。实际上，磷脂分子结构太复杂，即使在某些情况下能够在前生物条件下合成磷脂，仍难以相信是这样的分子形成了第一个原始细胞[2]。

相反，已经证明脂肪酸是前生物可利用的[4]，而且脂肪酸能够自发地自我装配成囊泡[12,13]。在一些可能的前生物条件下都能合成脂肪酸[14]，而且在 Murchison 陨石上也发现了脂肪酸[15]。与磷脂囊泡相比，脂肪酸囊泡具有不同的性质，主要取决于酰基链的长度及链中双键的数量。并且，脂肪酸自我装配成囊泡或是微团，取决于它们头部基团的电离程度（图 12.4），可根据表面活性参数适当调整。长链羧酸盐完全解离时可以聚集成微团（高 pH），而脂肪酸完全非解离时会从水溶液中分离成油滴（低 pH）。在中等 pH 时（通常为 7.5～9.5），脂肪酸及其共轭碱（羧酸盐形式）等量，能形成双层膜；图 12.4 中的氢键二聚体能够解释这种现象。这种分子间相互作用形成的二聚体，使得 pK 发生极大改变；在超分子膜中，长链脂肪酸的 pK 比短链水溶性单体羧酸（如乙酸）的 pK 约高4～5 个 pH 单位。

由这些结果及脂肪酸相对较高的水溶性（～mmol/L），得知脂肪酸具有复杂的相图，并通常比相应磷脂活性更高。脂肪酸要作为可能的早期原始细胞，这可能是对其有利的性质。在过去 10 年中，专业杂志中出现了越来越多的报道，用不同的技术来研究脂肪酸囊泡，关注它们的结构、形态、稳定性、理化性质和活性。有兴趣的读者可以查阅原始文献或关于这些研究的最近综述[4,5,16]。

在不同的脂肪酸中，油酸（即 Z，图 12.4）作为研究脂肪酸囊泡的模型系统，得到了特别关注。在下文中，将分析关键途径中对油酸/油酸盐囊泡，通常称为油酸盐囊泡的活性和转化的研究，它应该与原始细胞起源和进化有关。

除了脂肪酸，我们认为烷、寡异戊二烯磷酸盐、脂肪醇和单酰甘油也与前生物相关[5]。

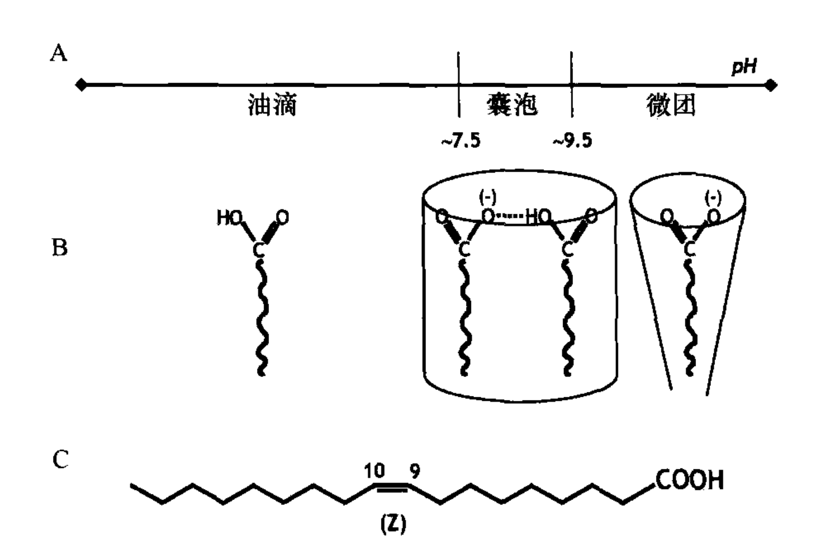


图 12.4 脂肪酸的自我装配过程。A. 通常脂肪酸会在低 pH 时从水溶液中分离成“油”，在中性 pH 时形成双层（囊泡），在高 pH（7.5～9.5 是典型例子）时形成微团。B. 人们把不同 pH 时的不同装配模式归因于分子的质子化状态不同，即在低 pH 时质子化，高 pH 时完全电离，pK 附近的中性 pH 时则形成 1∶1 的酸∶碱二聚物。C. 油酸的化学结构。

12.4 囊泡的活性和转化

囊泡是生物膜的唯一模型。为了得到细胞的更有意义的生物模型，膜内还需要有生化物质和与生物相关的反应。这个目标已经完成，在这里我们将回顾一些囊泡中反应的基础，主要是脂肪酸囊泡。

利用脂肪酸囊泡作为早期生物模型，有两个方面可以揭示这种人造结构单元是如何参与到生物起源研究中的，这两个方面是脂肪酸囊泡中基础生化反应的实现和脂肪酸囊泡的自我增殖。首先将分别讨论囊泡活性和转化这两个重要性质，之后我们将展示如何发现一条将它们的功能偶联的途径。需要注意的是，这里展示的模型系统受到自生说的启发，并且是第一次试图通过构建模型来具体理解活细胞的基本过程。

12.4.1 区域化反应

其主要观点是，在囊泡内腔中能够进行具有生化意义的简单和复杂反应。在前面章节中我们已经介绍过有关囊泡渗透性的基本概念。脂囊泡和脂肪酸囊泡对不同分子具有选择透过性。一方面，大分子和带电分子透过性较低，因而不能轻易穿过疏水膜，相反，小分子即使带电，也能以相当的速率透过双层屏障。此外，疏水和两亲分子能在膜上或膜内聚集。也就是说在某些情况下能通过外部添加前体，向囊泡中提供物质。例如，陷入囊泡的酶更易获得底物，底物渗透进囊

泡中，与陷入的酶反应，产物可能从囊泡中释放，也可能不释放，这取决于它的化学性质。

根据这种方式（可具有不同变异），人们研究了囊泡中的几种酶，如碳酸酐酶、α-胰凝乳蛋白酶、β-半乳糖苷酶等[17]。但是如何将酶插入囊泡内部呢？通常的方法是在溶有酶的溶液中准备囊泡。尽管在囊泡形成过程中，酶应该被动陷入，但在某些情况下，如碱性酶和带负电的表面活性剂时，静电力对组装产量和酶在囊泡中的位置影响很大。在囊泡形成后，剩余在外部环境中的酶可通过凝胶过滤、离心、透析或超滤除去。囊泡非常大时（$>10\mu m$），也可通过微注射技术将酶直接注入囊泡中。

早在 1990 年，为了建立与最小细胞结构相关的过程，Schmidli 等人就进行了最早的重要囊泡生化反应实验[18]：将负责合成卵磷脂的 4 种酶陷入到卵磷脂脂质体中，预计在脂质体中将生成卵磷脂，使脂质体从内部生长，甚至可能从内部分裂，这是典型的自生实验。这个实验部分成功了，但由于酶的商业供应中断，没有全部按预想完成。

为了模拟早期原始细胞起源的两个非常重要的过程，人们试图在脂肪酸囊泡中进行多 AMP（polyA）[19,20]的合成和 RNA 的自我复制[21]（图 12.5）。对第一种情况，将多核苷酸磷酸化酶（PNPase）装入油酸囊泡（pH＝9），并外部添加二磷酸腺苷（ADP）（图 12.5A），由于核苷酸不是完全不通透的，所以反应结束时从囊泡内分离出了 polyA，一旦聚合，产物实际上不能从囊泡内腔中逃脱。这个系统显示了聚合作用和半透性的结合如何产生有关生命“呈现”的结果，即形成包含大分子的原始细胞，这是一个含有功能性多聚物的功能性囊泡早期形成的模型。在远古时期，一种小催化剂可能起了 PNPase 的作用。注意，理论上囊泡膜也可能协助多聚反应，或者，囊泡中的反应可能导致形成一些分子，在之后的多聚作用中做催化剂。对第二种情况（图 12.5B），通过 Qβ-复制酶进行 RNA 链的复制[21]。酶和反应底物（RNA 和核苷酸）自发包裹在油酸囊泡内，进行 RNA 复制。囊泡内酶催化的 RNA 复制显然与功能性原始细胞的起源有关：它是早期区域化 RNA 复制的模型，要么是像 RNA 世界假说中一样被 RNA 自身催化，要么是被其他催化剂催化，如原始酶。

这些研究的关联（图 12.5B）也与囊泡的自发自我增殖相关，它们能够在区域化反应中共同进行，从而形成一个能够自我增殖的囊泡，且其内部能进行多聚反应和系统的自我复制，这些将在之后讨论。因此，这些模型即使不能解释第一个原始细胞是如何出现的，也能为理解区域化反应提供有力的实验途径。

在脂质体中也有可能进行 PCR 反应[22]；之后，核糖体也可能陷入囊泡内腔，从而第一次能在囊泡内合成多肽[23]。最后这个实验为随后一系列试图在脂质体中表达蛋白质的实验打开了大门。

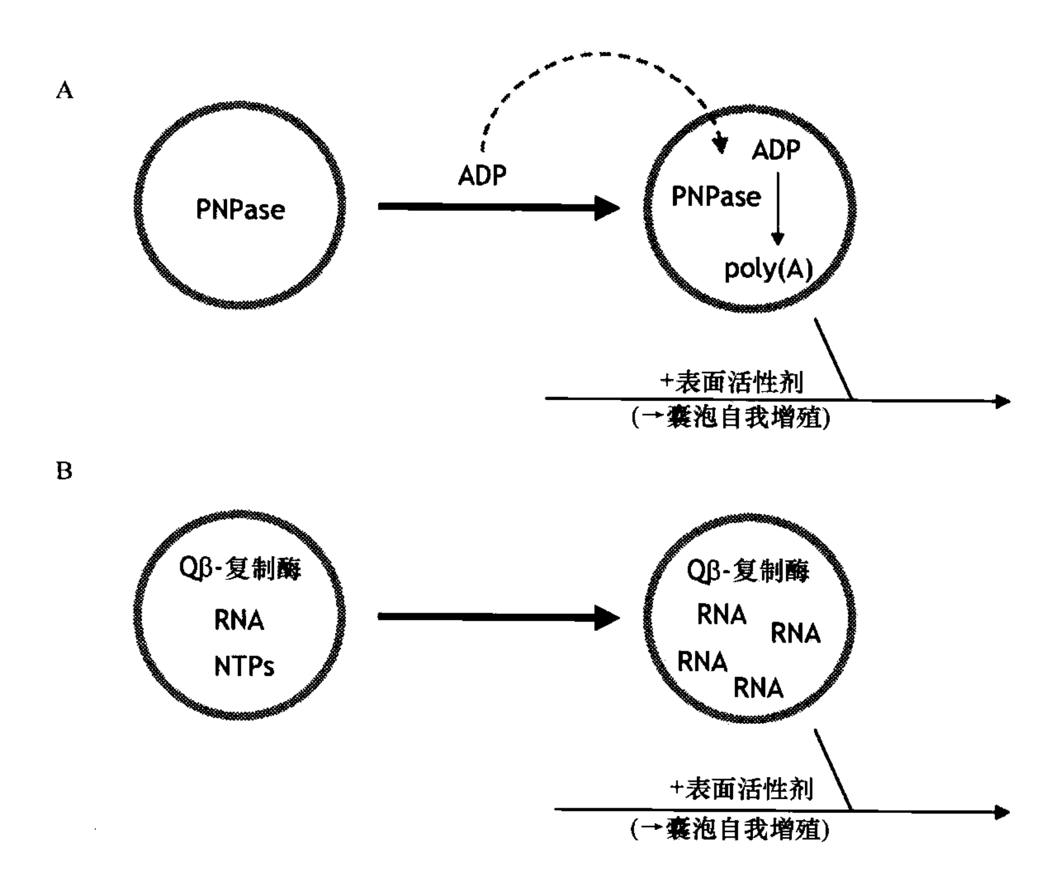


图 12.5 在脂肪酸囊泡中进行的两个重要反应。A. 多核苷酸磷酸化酶（PNPase）内陷在油酸囊泡中，然后外部添加 ADP；它缓慢过膜渗入囊泡，并在内部聚合生成 poly (A)。如箭头所示，这个过程在自我生成的囊泡中进行。B. 模板 RNA，Qβ-复制酶和所有小分子化合物都陷于油酸囊泡中，从而进行 RNA 复制，同时囊泡自我增殖。

如今，人们都承认，通过自动摄入基因和整套转录-翻译过程所需的生化机制，蛋白质就能在囊泡中表达。绿色荧光蛋白、T7 RNA 聚合酶和 α-溶血素已经在磷脂囊泡中成功得到表达。再次提醒，有兴趣的读者可以参考最近的综述，了解细节[2,24]。随后我们将讨论一些有关最小的细胞这个概念的实验。

12.4.2 囊泡的自我增殖

我们之前提到过囊泡能够进行“自我增殖”的能力，这显然非常重要，它提供了一个很接近生物细胞的模型。我们将在这里概述这个机制的基本原则，读者可以参考特定文献，进一步了解细节[2]。

像囊泡、微团和反相微团这样的超分子结构可以进行“自我增殖”过程，在这里，“自我增殖”这个术语非常重要，而模板分子实际进行的互补复制，则对小分子和核酸适用。与自我复制相反，自我增殖是通过生长—复制进行新结构的形成，基本不是像核酸复制那样具有分子水平的控制。分子自我复制，细胞则自我增殖。作为典型，囊泡可以生长，并分裂成两个不同大小的子细胞，因此初始

结构（囊泡）增殖形成两个类似但不相同的新囊泡。

20 世纪 90 年代，Luisi 及其同事对超分子结构的自我增殖进行了先驱性的探索，按年代顺序发现了反相微团、微团和囊泡的自我增殖[25]。

在 1994 年，人们利用脂肪酸囊泡和脂肪酸前体实现了囊泡的自我增殖（图 12.6）。实验方式明显受到自生说的启发，即实现了一个动力系统，按以下方式运行：一个具有边界的粒子（囊泡）摄入其生长所需的“养分”，在其边界内将它们转化成粒子的化合物（成膜分子），从而生长，达到稳定状态，然后分裂成两个或更多的新粒子。自生粒子前体到元素的转化必须在粒子内部进行，因为这是区分囊泡自我增殖和自发形成的关键内因。

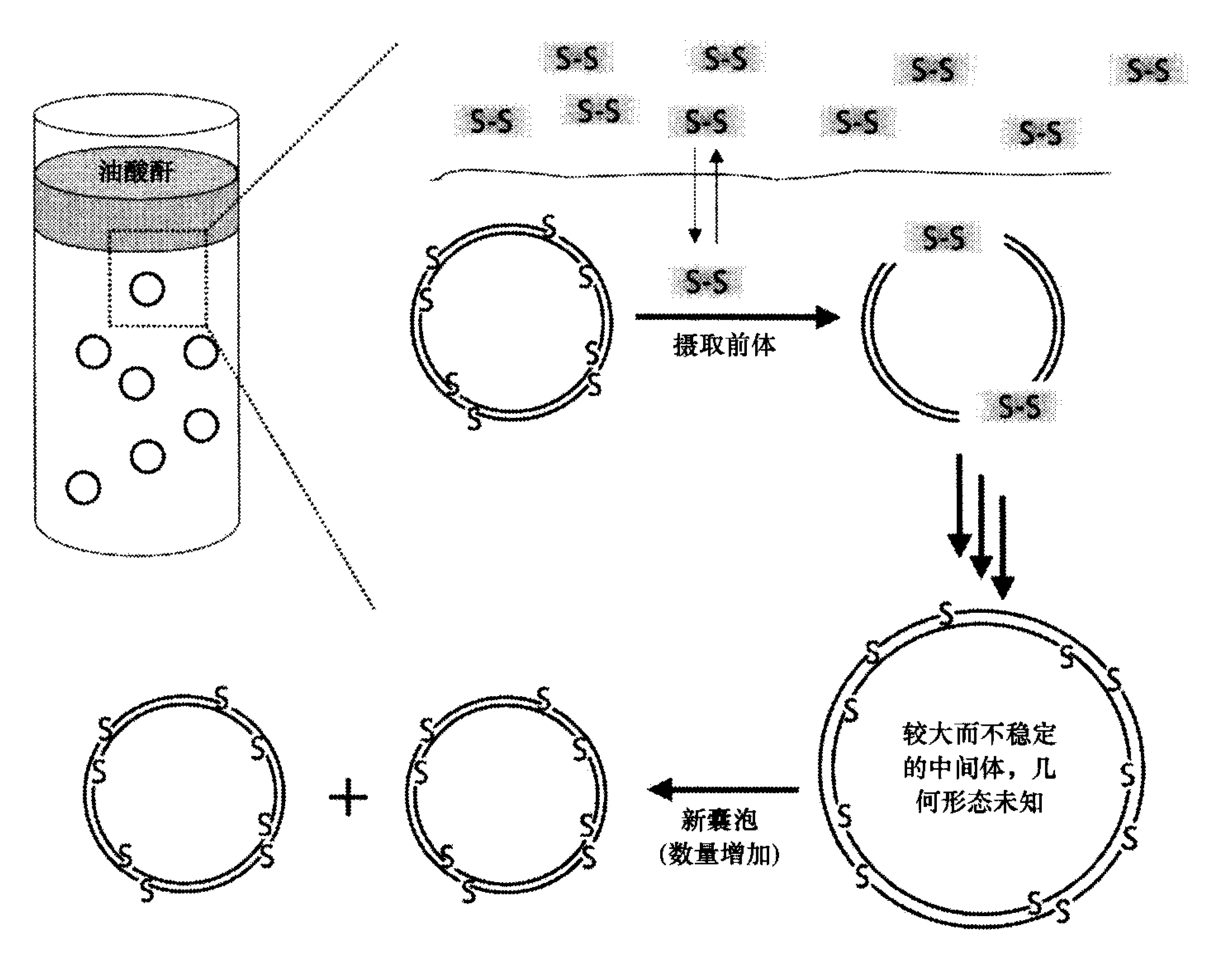


图 12.6　油酸囊泡的自我增殖。油酸酐为一种不溶于水的前体，覆于已含有油酸囊泡的碱性溶液上。油酸囊泡通过疏水双层溶解油酸酐分子，吸收它们，伴随水解增加其表面积。反应产物为两个油酸分子，增加了囊泡的表面积，并使囊泡生长（不一定是球形），最终使结构不稳定，分裂成两个或更多个小囊泡。整个过程使囊泡大量自催化生长。若起初油酸囊泡不存在，那么在油酸酐分子水解之后，也会自发形成囊泡，并在经过一段时间之后催化整个过程。

实验方面，可以在含有油酸囊泡的碱性溶液中，放上油酸酐（前体）层，来实现这个过程。在溶液的分界面上，油酸酐以较慢的速率水解。同时，油酸囊泡

摄入其他油酸酐分子，这是因为它们可以在双层膜中进行溶液化作用。一旦进入膜中，油酸酐就被水解，转化成两个油酸分子，相应地膜表面就增加了。通过膜前体的加入，转化成膜本身的结构单元，油酸囊泡从而生长，最终产生新囊泡（其机制仍了解得很少）。

生长—分裂过程的化学计量学公式可被简要概括为如下形式：

$$V+nS\longrightarrow 2V$$

其中，V是一个囊泡；S是成膜化合物（为了简化，S也可作为前体）。因此，这个过程是自动催化的，因为囊泡产生得越多，摄入的前体就越多，就能产生更多囊泡等。在实验中，自动催化过程的特征是反曲线，而油酸囊泡自我增殖实际上就是这样（图12.7）。还需注意，如果预先形成的囊泡不包括在实验设定内，它们也能随着酐自发水解形成，其结果是表面活性前体形成囊泡的整个过程成为自发途径，不需要预先形成的结构。进入自我增殖循环的驱动力来源于酐水解和脂肪酸自我装配产生的热力。

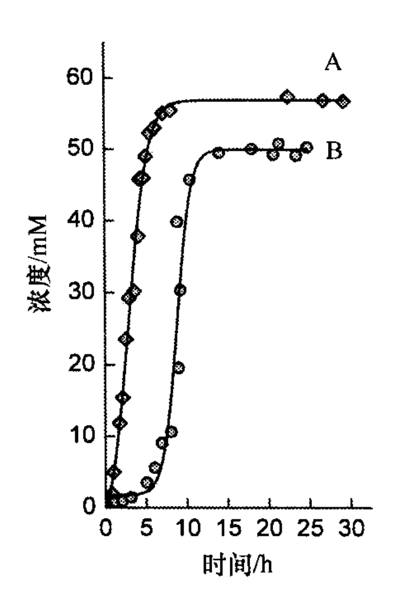


图12.7　图12.6系统的动力学特点。若水相中已含有油酸囊泡，油酸酐会迅速水解（曲线A）。若刚开始时没有油酸囊泡，它们会随着油酸酐碱性水解，达到临界聚合浓度时形成。S形曲线（曲线B）清楚显示了自动催化机制，即囊泡催化形成其他囊泡。

油酸囊泡自我装配的第二个例子体现了脂肪酸依赖pH的装配特性。实际上，因为高pH时油酸分子自我装配成微团，而在中性pH（8.5）条件下它们则形成双层囊泡，因此，向含油酸囊泡的缓冲溶液中加入油酸微团也是补充囊泡的途径之一，可用来观察它们的生长和分裂，即自我增殖。事实上很多研究者目前都采用这个方法，它没有油酸酐途径中涉及双相系统的不利之处。

油酸微团和油酸囊泡都是宏观单相系统，因此利用分光技术很容易追踪囊泡的自我增殖过程。当向pH 8.5的缓冲溶液中加入油酸微团（与单体形式等量）时，它们能重排组装模式，自发形成囊泡。但是，若在这种溶液中存在预先形成的油酸囊泡，则可能有另一条途径，由事先形成的囊泡摄入油酸（以单体或微团的形式），正如之前油酸囊泡和油酸酐的情况一样。在第二种途径中，事先形成的囊泡会生长、分裂，开始进行自我增殖。这两条竞争途径可以利用水溶性标记物，如铁蛋白分子，标记事先形成的囊泡（图12.8）。通过测量添加油酸微团前后囊泡中铁蛋白分子的分布情况，表明囊泡确实可以生长，并通过囊泡的分裂，生成少量但显著的含有铁蛋白的囊泡，每个囊泡中铁蛋白数量会减少，这是铁蛋白分子被分配到“子”囊泡的结果。

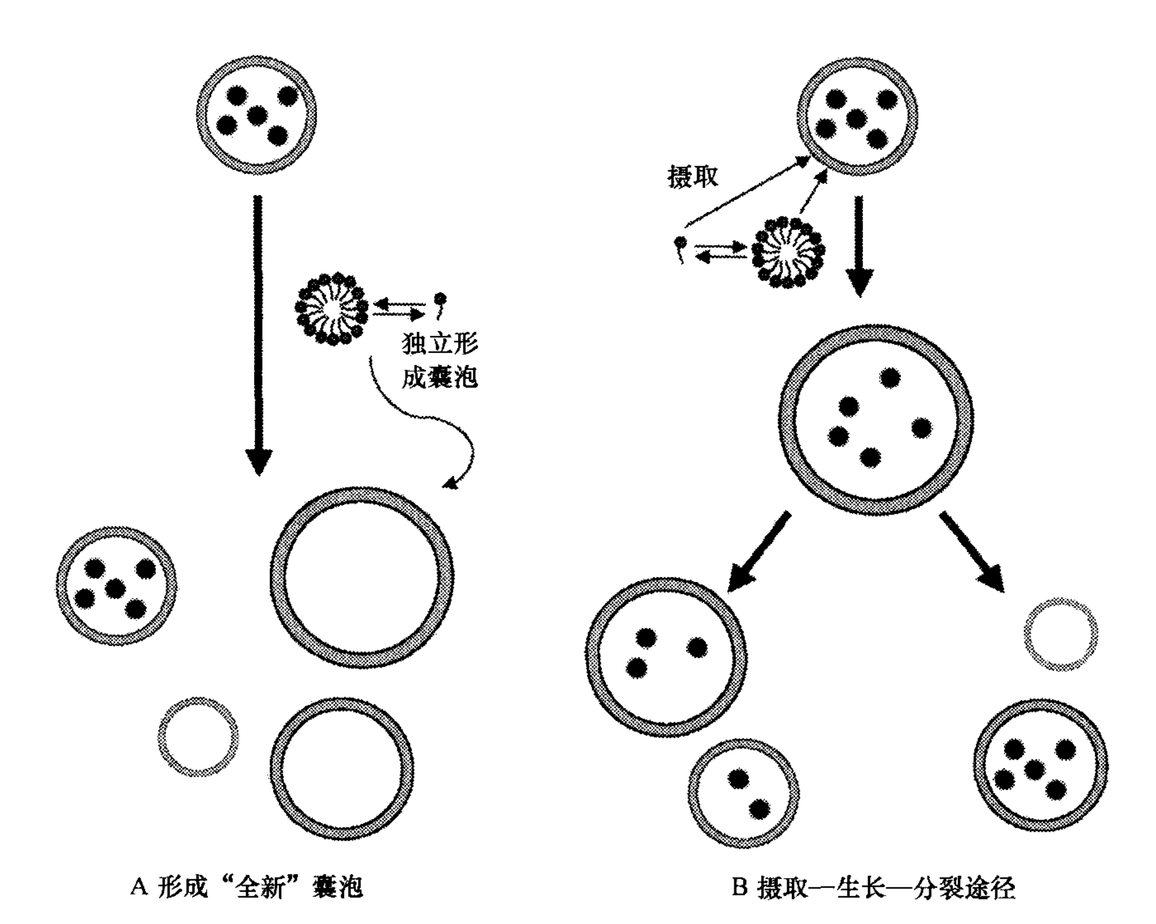


图 12.8　利用铁蛋白作为水溶性指标，检测油酸囊泡的形成。第一种情况（A），油酸微团独立形成全新油酸囊泡（实验中利用卵磷脂囊泡作为预先存在的囊泡）。其机制中，油酸微团或油酸单体与预先存在的囊泡间没有相互作用，因此，后者的标记物没有被稀释。第二种情况（B），油酸微团或油酸单体被预先存在的卵磷脂囊泡吸收，作为新的成膜分子生长，形成膜。在这个初始阶段之后，囊泡变得不稳定（生长也可能不是球形的），分裂成两个或更多“子”囊泡，因此，在新囊泡中探针可能不符合统计学分布，也可能符合。每个囊泡中平均铁蛋白分子数量的下降意味着存在生长—分裂途径（囊泡、微团、脂肪酸和铁蛋白分子没有按比例绘制；圆点代表铁蛋白）。

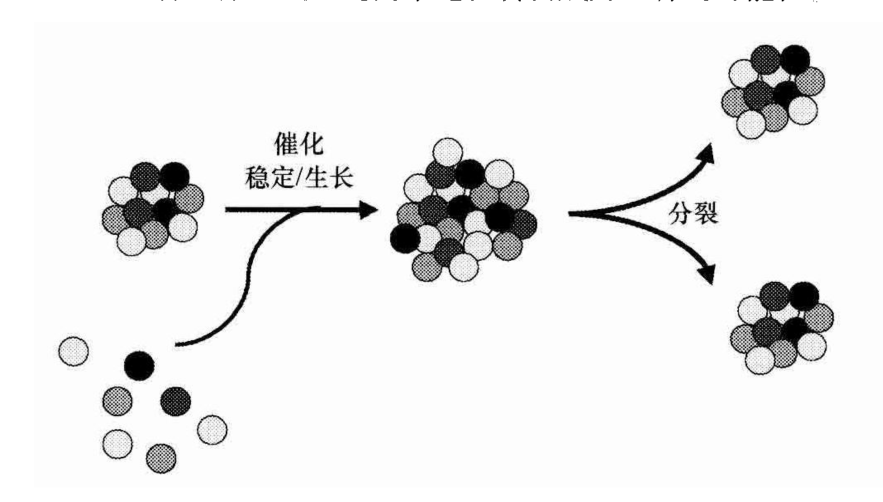
对脂肪酸的自我增殖，现代研究揭示了一系列有趣的特征，在此不做详述，比如“母体”效应，是由事先形成的囊泡产生的一类“模板”效应，导致形成大小一定的囊泡群[26]；不同大小的囊泡摄入油酸微团时的竞争[27]；渗透性“强”与“松弛”的油酸囊泡间的竞争[28]。

专题 12.1 描述了 Doron Lancet 的原创性贡献，将囊泡自我增殖和可遗传的“组成”信息结合起来。

专题 12.1　脂类“合成体”（composome）及它们作为前生物复制子的作用

超分子聚合体，如脂类和脂肪酸囊泡，在生命起源中的作用，不仅仅和它们能为生化反应发展提供结构单元有关。以色列 Rehovot Weizmann 研究所的 Doron Lancet 等提出了一种有可能性，然而也很奇特的分级自动催化复制域（GARD），这是一种无基因信息

复制方式[29,30]。他们提出，在前生物进化的最初阶段，囊泡这样由几种分子化合物组成的超分子聚合体，能稳定生长和分裂，产生与其母体有相同成分的新的聚合体，选择性吸附和环境中化合物的转化有利于稳定生长。以这种方式，一种具有特定组成和组织方式的聚合体就通过聚合体的产生传递了“信息”。精细的计算机模拟已显示了这种“合成体”的存在，及由一种合成体形式经过简单进化转化成另一种的可能性。



图专题 12.1.1 “合 成体”，能够稳定生长，且分裂时能保留信息的特殊组成状态。

12.4.3 核和壳转化的功能性结合方式

我们已知：①简单和复杂的生化反应（包括自我增殖）可通过囊泡进行；②通过由外部添加合适的表面活性前体，囊泡能自发地自我增殖。由这些来自于实验的考虑出发，可以想象，是否有可能将这两个过程进一步结合，建立一个自生自我复制细胞的简单模型。

在理论上，很容易看到答案：将能自我增殖的代谢网络封闭在囊泡内，网络中的一些过程应该能够由合适前体产生成膜化合物，系统在刚开始时构建简单，然后变复杂（图 12.9）。囊泡生长的方式，可以基于循环网络的分子内在化，随反应的进行产生越来越多的膜分子或膜表面。在囊泡自我增殖的过程中，可能是不稳定状态的瓦解产生两个或更多“子”囊泡，每个含有部分内在化组分。如果内在化组分的数量很多，那么每个子囊泡包含所有组分的可能性就高，“活性”囊泡的数量也会增加。反之，由于这个过程是不被控制的，内在化分子的分裂可能不均衡，而使一些子囊泡缺少某些组分，不能维持代谢网络。

图 12.9 中描述的系统是自生的，根据 Maturana 和 Varela 的最初看法[7]，它甚至可能被认为是活的；但最近的研究使我们要对这个观点重新斟酌[8,31]。回到实验方面，人们可能会问如何能在实验室中实现图 12.9 中展示的系统。

这个方向的第一次尝试是由 Zepik 等实现的[32]，通过简化内部反应网络，建立了图 12.9 中的模型，只含有两个反应：S 这种成膜分子的合成和降解。两个

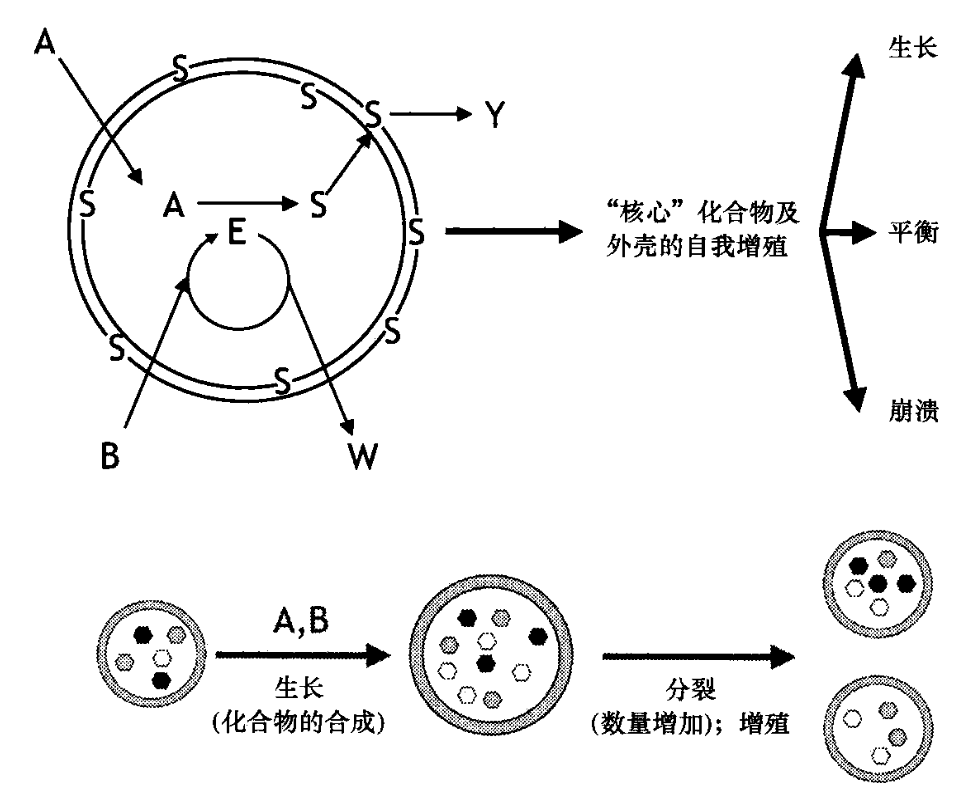


图 12.9　自生原始细胞的自我增殖。囊泡内腔中有一个代谢网络。代谢网络"E"（由一种或多种酶组成）催化前体 A 形成成膜化合物 S，A 最初存在于环境中，通过渗透作用过膜。S 会衰退为 Y。这里用"E"表示的整个代谢网络，本身通过使用一系列前体"B"，产生废物"W"，从而能在囊泡中再生，这就是核-壳自我增殖作用。注意：内部化合物和壳上化合物 S 的形成是功能相关的，这是因为内部网络的产物催化形成了 S。基于构建（合成代谢）和毁灭（分解代谢）过程的平衡，原始细胞能够经历生长、崩坍，或当过程平衡时进入一种自我平衡的状态。当原始细胞合成其化合物时，它就能生长、分裂，其内部化合物统计上也能分配到"子"细胞中。由于不能控制囊泡分裂，因而某些新囊泡可能会缺乏关键组分，从而进入"死亡"。

反应的化学过程（S 从酐的水解形成和 S 氧化耗损）可以调节速率，使人们能够观察图 12.9 右手边的三个不同状态。通过适当调节合成和分解过程，达到自我平衡。

从生命起源的观点来看，很重要的一点是需要证明，简单化学系统能够以图 12.9中的方式进化、自我装配，并自我组织。但是，以现有的知识水平，这个目的似乎很难实现，尤其是需要在系统中建立一定程度的复杂性和（或）精确性时。并且，如果想要重建由简单分子到第一个功能性大分子（原始酶或核酶），到第一个"代谢"循环，以及自我复制或自我增殖的途径，人们就会发现基础知识的不足。可以说，尽管投入了大量实验和理论研究，但是几乎所有的关键转变都还没有确知。

关于功能性大分子的自发产生（在目前对生命起源的研究中，这是最大的挑

战之一)，有一个可能机制，在前生物可能存在的自我装配的囊泡，如脂肪酸囊泡中，产生封闭代谢的初始。从这个阶段开始，内在化反应网络将能够自我维持和自我复制，然后产生遗传密码，再之后是编码自我复制和核-壳的自我增殖(图 12.10)。在这个原始细胞“进化”的途径中，与环境的相互作用是基础，因为内在化代谢需以某种方式和外部介质的组成相联系。囊泡自我增殖由外部供应或内部合成的表面活性剂介导，能够起到原始细胞扩增的作用，它起初是随机的，逐渐产生差异和原始细胞间的竞争，随后受到内在化代谢的控制。其他现象，如融合、生长-分裂、竞争、选择性过程和囊泡的整个反应情况都利于超分子系统逃离无生命物质的世界。这个假设在很大程度上仍属推测，正关注于区域化反应网络个体和群体的概念，以及它们可能的进化和复杂性的逐步增加。

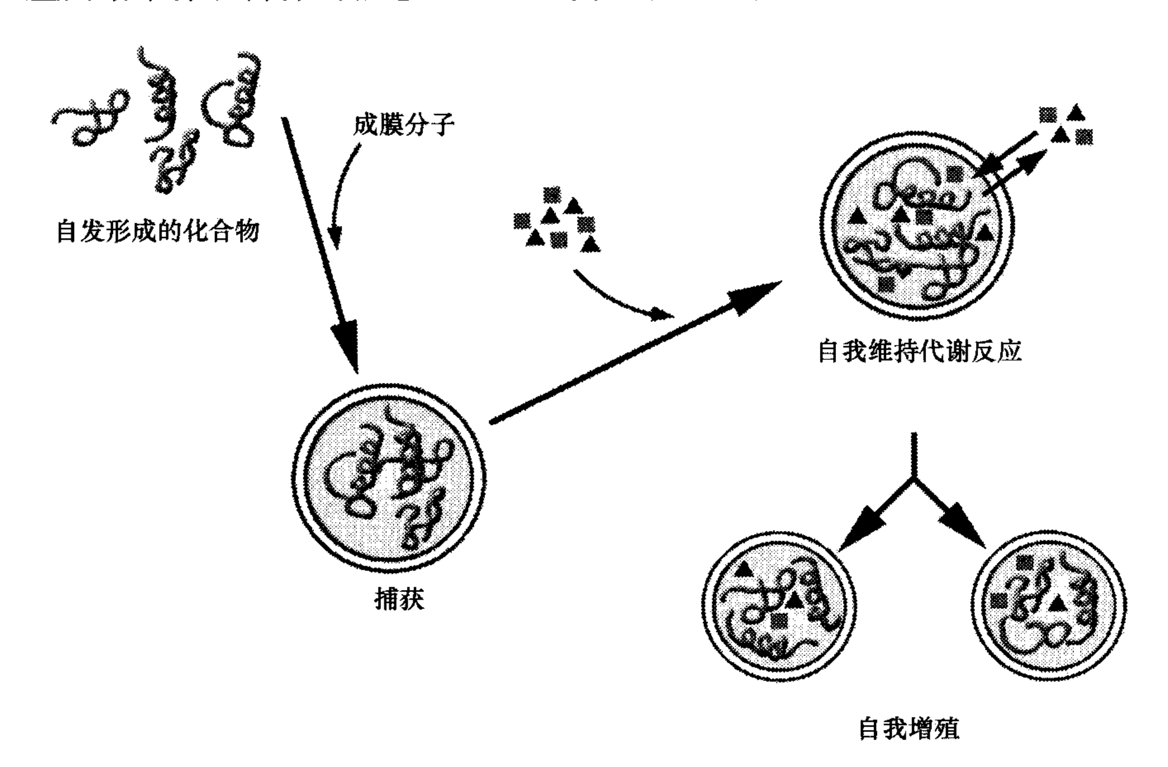


图 12.10 原始细胞产生过程的臆测图解。功能性生物多聚物和表面活性剂自我装配成细胞样结构，即囊泡。由这个初始阶段开始，系统必须能够自我维持、自我增殖，并通过逐步的化学进化直接合成膜。这种转变的许多步骤都还未清楚，尤其是第一个功能性大分子的起源仍然未知。

为了理解图 12.10 中的一些基本转变，近年来有一些实验和理论的探索，尤其重要的是解答了大量关于最早细胞性质和功能的问题，它们不具有成熟的基因、酶、适当的代谢调节，反馈环和控制功能。基于这个初期细胞的“历史性”面目与理论上使一个系统“活起来”所需的最少的化合物和功能上完美的互补，我们能构建这样一个系统吗?

在可信、具体、完善的实验（即使不容易实现）的框架下，我们可以“限

制”自己来研究半合成结构，利用外部酶（基因）和合成单元来构建“最小的细胞”，即类似于细胞的结构，具有图 12.9 中展示的活的性质。当然，在没有已知原始（特异性更小）的“酶”时，先利用进化过的酶和基因，也很适合原始细胞的设计和构建。

这个目标的实现将代表，甚至超越早期活细胞重建的历史意义，它是生命作为区域化的化合物/反应网络的调节和总体性质的出现和产生的原理论证。

最近的研究进展（综述见参考文献 24）表明，实验者们已经能够获得具有复杂内在化反应的囊泡，如蛋白质表达，反应为单一步骤或级联反应，并可以在几小时到几天的时间内进行观察。

总之，关键仍然是内在化代谢反应（意思是功能性大分子的形成和复制）和整个囊泡增殖的结合。这种结合必须是功能性的，即这两个过程必须通过反应或反应网络，或这两个途径共有的分子结合。之前已经论证过，内在化反应和囊泡的自我增殖可以同时发生，如 Qβ-复制酶进行 RNA 复制和囊泡的自我增殖，但是在这个例子中，这两个过程只是同时进行，在核和壳的复制中并没有建立起功能联系。

图 12.11 给出了一条可行途径，囊泡在有催化剂（一个酶或一套酶）存在时合成膜化合物，也就是通过内在反应网络进行自我增殖。原则上，这可以用体外功能蛋白表达和前面提到的囊泡生长-分裂过程的结合来实现。最终，产生酶所需要的所有化合物都需要得到复制，并在子囊泡中均匀分配。这就会构建出一个半合成的最小细胞[2,24]。

这样的最小细胞远远称不上完美，但能丰富我们对早期细胞生物合成和生物特性的理解，它们还需要底物的渗透性、内在化反应动力、多余材料的清除、能量、同时进行的核-壳复制，以及与环境和其他细胞的相互作用。图 12.9 和图 12.11 中那样的构建目前尚未能实现。但我们相信，在实现这个目标的过程中，能够洞悉细胞生命起源的一些重要事件。这种“瘸腿”细胞，缺乏许多先进功能，但可能是早期细胞的重要模型，并可以对它进行实验检测和验证。

12.5 结语

结构单元，尤其是囊泡，在生命起源中起了奠基作用。我们已经知道，除了为分子反应网络的发生和维持提供空间范围，结构单元的形成更奠定了独立个体的出现，在达到自生条件时，实现了自我维持、自我边界和内部平衡。此外，将自我组织的反应网络限制在囊泡内，在“系统”和环境间建立了清晰的界线。一些作者已经从结构单元的理论观点意识到这些，尽管实验方法最近才被引入。最清晰的立场之一是 Harold J. Morowitz 提出的，他宣称[1]：封闭囊泡的形成是生物发生过程中的主要事件，它代表了超分子整体结构的起源，它是三相系统，由极性的内部、非极性膜核心和极性的外部环境构成。在细胞发生过程中，怎样看

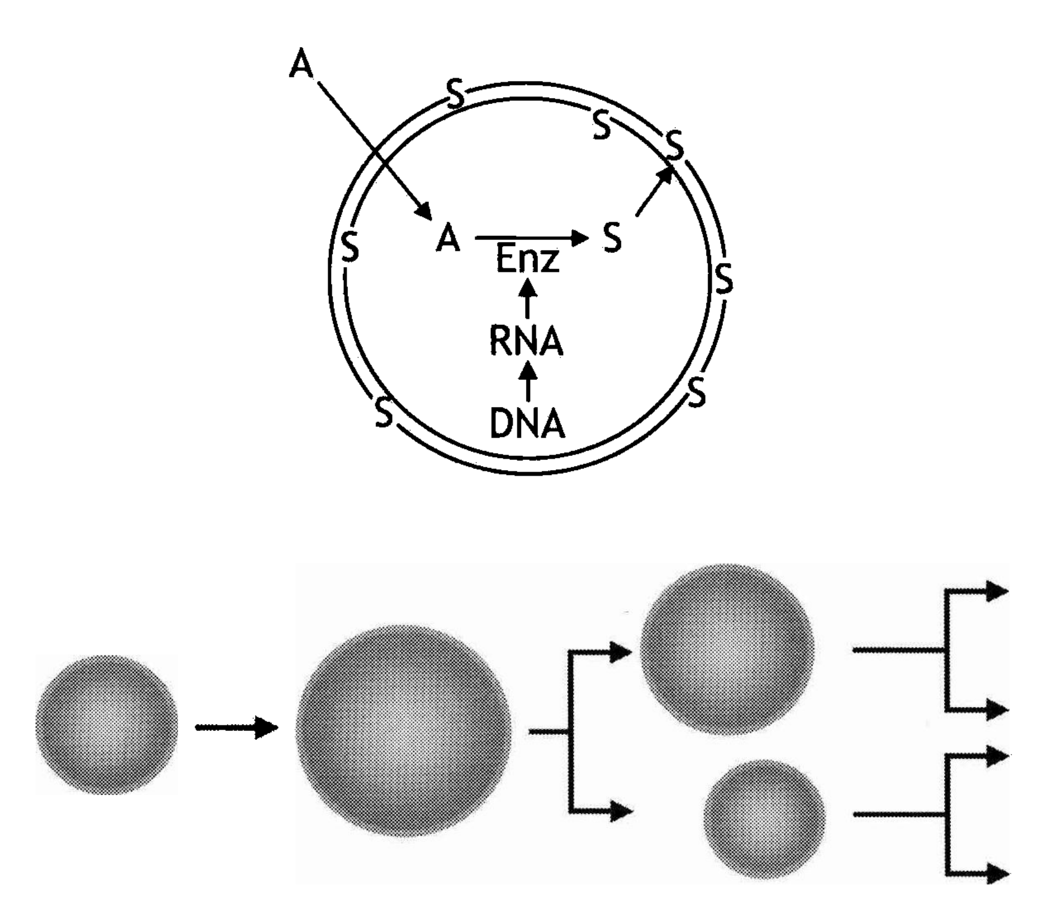


图 12.11 获得半合成最小细胞的实验方法。酶、基因、转录/翻译机制都内陷于囊泡中。原地合成的酶能够促进合适的前体产生成膜化合物 S，使囊泡生长和分裂。若不能生成内部化合物，系统就会在一系列的分裂之后死亡，这些分裂是由于某些囊泡中的网络元素过度“稀释”造成的。

待封闭囊泡的重要性都是不过分的。在最普遍的意义上，这个事件建立了生命体与环境的本质对立。通过限制性渗透屏障，封闭带来了内部和外部之间的物理隔离。若没有这层屏障，细胞的基本定义都无从想象。

人们不知道第一个活细胞是如何起源的，但清楚的是，没有结构单元就没有生命，因为生命是一种呈现的整体、来源于自我组织的复杂分子系统。生命并不是体现在单个分子中，如 DNA 或核酶，而是来自于自主产生的自动催化和循环反应的结合。即使不能解释细胞生命如何在地球上起源，自生说仍为描述生命提供了强有力的观点。

实验显示，一些反应，包括具有显著生化意义的反应，可以在脂类或脂肪酸囊泡狭小的内部水环境中重建，通过成膜分子前体的结合，这些囊泡能够生长和分裂。这些振奋人心的发现推动了现今对实现半合成结构的研究，用囊泡封闭复杂的生化途径，从基因表达到囊泡渗透性控制，再到内部调节和脂类的生物合成，都是为了建立早期细胞的模型，令这些细胞具有生命所需的最少和最必须的功能，这个研究领域称为“最小细胞”领域，对其他学科也充满了吸引力，如合成生物学中的

构建方法。很明显，虽然半合成的最小细胞不能回答生命起源这个历史性问题，但是这些研究会使我们对细胞生命基本原理有更进一步的理解和验证。

回到这一章的主要论点——原始细胞的起源，我们必须承认对它们形成的机制目前了解得很少。我们相信，结构单元的形成是一个早期阶段，而囊泡的多种有利性质都强调了它们在从非生命到生命物质的转变中的重要性与必要性。他们的有利性能包括反应物的共区域化、对外部分子寄生物和抑制剂的防护、潜在的表面催化剂、膜中疏水色素的积累、电渗透和化学梯度的维持、大分子的储存、个体性的出现，以及最终形成自我维持的生化组织系统。不容置疑，脂肪酸囊泡是目前早期结构单元最合理的模型：除了在可能前生物条件下，其脂肪酸单体能够被合成，这种囊泡还能够由单体自动装配形成，生成半渗透性的边界，主导生化反应，自发生长分裂，从而实现自我增殖。

专题 12.2　其他生命起源学说中的结构单元

“结构途径”作为一种范式，已被应用于一系列实验和理论研究中。不过，明显的是，仅有结构单元是不足以产生第一个细胞的。不同的学者以各自的偏好和侧重解读了结构单元的意义，其最重要的贡献就是超越了代谢优先还是基因优先的争论。在本段中，我们将简述其中一些学者在他们的生命起源学说中是如何阐释结构单元的存在和功能的（见参考文献 6）。

Manfred Eigen 持基因优先观点，他提出，一些“协同”信息载体分子（RNA）是稳定共存的，且它们能自发组织成成熟的催化循环，即“超循环”[33]。这种途径中，一些复制子彼此协同作用，而不是相互竞争。在最初的构想中，有一个由 RNA 复制子组成的超循环，它能翻译成复制子酶，同时也存在核酶超循环。Eigen 的超循环克服了复制子错误阈限的限制，从而解决了形成高功能分子系列的问题（Eigen 悖论）。超循环的一个问题是，它不是独立个体，而细菌是独立个体。相反，超循环是相互作用的分子的集合，受分子寄生物、短连接和突变的影响。要稳定超循环，促进其适应性，有人谨慎提出，超循环可能插入到结构单元中[34]，以这种方式产生独立性，当超循环过程和结构单元复制结合时，就会产生可遗传的遗传信息。同时，结构单元保护超循环不受寄生物和其他循环的影响，并有利突变[35,36]。类似地，结构化超循环能够在原始细胞间带来大量竞争，通过自然选择产生高效的酶，从而产生更长的基因[6]。

Freeman Dyson 主张[37]由类蛋白质分子催化的有机反应网络在核酸产生和复制之前就出现了。他的理论受 Oparin 的代谢优先/凝聚起源论影响很深。根据 Dyson 的说法，包含类蛋白质催化剂代谢系统的结构单元和原始细胞是首要且必需的，这样的原始“细胞”，主要来源于 Oparin 的凝聚论，可以通过吸收环境中的基础材料来生长，并通过分裂进行增殖，因而组成性（也是功能性）的遗传基本模式变为可能。随后由这样的代谢产物（或作为细胞侵入物）产生核酸，核酸首先是代谢的寄生物，随后是共生物，最后完全整合为细胞组分。Dyson 模型中，结构单元是必要元素，封装了能够产生原始代谢反应的特定化合物。子结构遗传了母体的代谢物、原始酶和代谢途径的系统，即代谢反应的组成

及整个系统的框架。虽然 Dyson 的模型缺少特定的化学上的考虑，尤其是没有确定边界分子的性质，但它明显依赖结构单元和结构化反应来产生独立性，并设想了不一定依赖核酸的遗传途径。Dyson 的模型把内部平衡看做活细胞的基本性质。

另一个重要的理论模型是几年前由 Stuart Kauffman 构建的[38]。它根据复杂系统自我组织的范式，提出一些系统能够自发达到“有序”状态，摒弃初始的“无序”状态。要完成这种无序到有序的转变，系统必须是热力学开放系统，从而将增加的熵散入环境，减少系统的熵；这样系统就能在不违反热力学第二定律的情况下，维持其远离平衡的状态。这种观点来源于 Ilya Prigogine 的耗散系统[39]。Kauffman 声称，自我组织系统的自发产生是生命起源的关键因素。由此观点，他建立了一个自生反应网络模型，由催化剂及相互作用的生物多聚物组成。单独的分子不能自我复制，而系统，即相互作用的催化剂和反应网络，需要以整体为单位进行自我增殖。根据这一观点，一旦分子复杂性超过了一个关键水平，集体性就产生了。从化学角度看，催化性生物多聚物是核酸和多肽，不过这个理论并不包括自我复制的 RNA。结构单元从两方面支持 Kauffman 的理论，即使这个理论并不关注结构单元途径。首先，为了增加生物多聚物的局部浓度使反应进行，它需要一种限制，其次是引入进化/选择途径带来的需要。新的偶然突变的多聚物也可能加入到最初的自催化系统中，使系统走向不同的方向。如果自催化系统被限制到一些原始细胞（如 Oparin 的凝聚体、Fox 的微球体和脂质体）中，那么分裂过程就会将不同多聚物随机带到不同子细胞中，尤其是母细胞中多聚物的拷贝量很少的时候。这将使得自动催化系统通过大量原始细胞分裂得到进化。Kauffman 同时提出，除了脂质体的限制功能，它们通过渗透效应和熵效应，尤其能够促进生物多聚物的形成[38]。

Kunihiko Kaneko 提出了“少数控制”理论，为基因优先或是代谢优先的问题提供了新的视角，即前者缺乏应对寄生物的稳定性，而后者有遗传上的问题。在他的理论中，结构单元（原始细胞）的存在是最基本的需要；同时也需要分子和原始细胞两个水平的复制，认为结构单元对遗传信息的起源十分重要。由于复制缓慢，少数分子的浓度很低。但是，由于参与原始细胞整体复制所需要的共同催化作用，它们实际上控制了原始细胞的动力。少数分子必须被保留，而它们的数量又总在波动，因而极大影响了原始细胞的分裂时间，这必须得到抑制。关于少数控制理论的详细概念，有兴趣的读者可参阅 Kaneko 的专著[40]。

参考文献

扩展阅读

1. Morowitz H. Beginnings of cellular life. Metabolism recapitulates biogenesis. Yale University Press, New Haven, 1992.
   - 文章重点是结构单元途径及代谢的组织、网络复杂性和能源。
2. Luisi PL. The emergence of life. From chemical origins to synthetic biology. Cambridge University Press, Cambridge, 2006.
   - 该书描述了有关生命起源的理论及试验问题与细胞模型。
3. Luisi PL. Autopoiesis: A review and a reappraisal. Naturwissenschaften 2003; 90:49-59.
   - 文章介绍了自生说，包括它的历史、细胞生命的逻辑及生命的呈现、自生。
4. Monnard PA, Deamer DW. Membrane self-assembly processes: Steps toward the first cellular life. Anatomical Records 2002; 268:196-207.
   - 有关磷脂及脂肪酸囊泡在生命起源中的角色的综述。
5. Walde P. Surfactant assemblies and their various possible roles for the origin(s) of life. Orig Life Evol Biosph 2006; 36:109-150.
   - 该综述收集了大量有关表面活性分子的装配的数据，不单是囊泡，亦包括微胞团和反向微胞团。

专题文献

6. Fry I. The emergence of life on Earth. London:Free Association Books, 1999.
7. Maturana HR, Varela FJ. Autopoiesis and cognition: The realization of the living. Dordrecht:Reidel, 1980.
8. Bitbol M, Luisi PL. Autopoiesis with or without cognition: defining life at its edge. JR Soc Interface 2004; 1:99-107.
9. Damiano L. Unità in dialogo. Mondadori Milano, 2007.
10. Oparin AI. Origin of life on Earth. Edinburgh:Oliver and Boyd, 1957.
11. Fox SW. Proteinoid experiments and evolutionary theory. In: Ho MW, Saunders PT, eds., Beyond Neo-Darwinism. New York:Academic Press 1985:15-60.
12. Gebicki JM, Hicks M. Ufasomes are stable particles surrounded by unsaturated fatty acid membranes. Nature 1973; 243:232-234.
13. Gebicki JM, Hicks M. Preparation and properties of vesicles enclosed by fatty acid membranes. Chem Phys Lipids 1976; 16:142-160.
14. Rushdi AI, Simoneit BRT. Lipid formation by aqueous fischer-tropsch-type synthesis over a temperature range of 100-400°C. Orig Life Evol Biosph 2001; 31:103-118.
15. Lawless JG, Yuen GU. Quantitation of monocarboxylic acids in the murchison carbonaceous meteorite. Nature 1979; 282:431-454.
16. Walde P, Namani T, Morigaki K et al. Formation and properties of fatty acid vesicles (liposomes). In: Gregoriadis G ed., Liposome Technology, 3rd edition, Vol. I. New York:Informa Healthcare, 2006:1-19.
17. Walde P, Ishikawa S. Enzymes inside lipid vesicles: preparation, reactivity and applications. Biomol Bioengin 2001; 18:143-177.
18. Schmidli PK, Schurtenberger P, Luisi PL. Liposome-mediated enzymatic synthesis of phosphatidylcholine as an approach to self-replicating liposomes, J Am Chem Soc 1991; 113:8127-8130.
19. Chakrabarti AC, Breaker RR, Joyce GF et al. Production of RNA by a polymerase protein encapsulated within phospholipid vesicles. J Mol Evol 1994; 39:555-559.
20. Walde P, Goto A, Monnard PA et al. Oparin's reactions revisited: enzymatic synthesis of poly (adenylic acid) in micelles and self-reproducing vesicles. J Am Chem Soc 1994; 116:7541-7544.
21. Oberholzer T, Wick R, Luisi PL et al. Enzymatic RNA replication in self-reproducing vesicles: an approach to a minimal cell. Biochem Biophys Res Commun 1995; 207:250-257.
22. Oberholzer T, Albrizio M, Luisi PL. Polymerase chain reaction in liposomes. Chem and Biol 1995; 2:677-682.
23. Oberholzer T, Nierhaus KH, Luisi PL. Protein expression in liposomes. Biochem Biophys Res Comm 1999; 261:238-241.
24. Luisi PL, Ferri F, Stano P. Approaches to semi-synthetic minimal cells: a review. Naturwissenschaften 2006; 93:1-13.
25. Luisi PL. Self-reproduction of micelles and vesicles: Models for the mechanisms of life from the perspective of compartmented chemistry. Advances Chem Physics XCII, J Wiley and Sons Inc 1996:425-438.
26. Blochliger E, Blocher M, Walde P et al. Matrix effect in the size distribution of fatty acid vesicles. J Phys Chem 1998; 102:10383-10390.
27. Cheng Z, Luisi PL. Coexistence and mutual competition of vesicles with different size distributions. J Phys Chem B 2003; 107:10940-10945.
28. Chen IA, Roberts RW, Szostak JW. The emergence of competition between model protocells. Science 2004; 305:1474-1476.
29. Segre D, Ben Eli D, Lancet D. Compositional genomes: prebiotic information transfer in mutually catalytic noncovalent assemblies. Proc Natl Acad Sci USA 2000; 97:4112-4117.
30. Segre D, Ben Eli D, Deamer D et al. The lipid world. Orig Life Evol Biosph 2001; 31:119-145.
31. Bourgine P, Stewart J. Autopoiesis and cognition. Artificial Life 2004; 10:327-345.
32. Zepik HH, Bloechliger E, Luisi PL. A chemical model of homeostasis. Angew Chemie Int Ed 2001; 40:199-202.
33. Eigen M, Schuster P. The hypercycle: A principle of natural self-organization. Part A. emergence of the hypercycle. Naturwissenschaften 1977; 64:541-565.
34. Eigen M, Schuster P, Gardiner W et al. The origin of genetic information. Scientific American 1981; 244:78-94.
35. Eigen M. Stufen zum Leben: die frühe evolution in visier der molekularbiologie. München:Piper, 1987.
36. Maynard Smith J, Szathmáry E. The major transitions in evolution. Oxford:Oxford University Press, 1995.
37. Dyson FJ. The origins of life. Cambridge:Cambridge University Press, 1985.
38. Kauffman SA. The origins of order. Self-organization and selection in evolution. New York:Oxford University Press, 1993.
39. Nicolis G, Prigogine I. Self-organization in non-equilibrium systems. New York:John Wiley, 1977.
40. Kaneko K. Life: An introduction to complex systems biology. Berlin:Springer, 2006.

13 断裂基因、始祖基因

Massimo Di Giulio

13.1 引言：基因的起源

一想到基因的起源，我们几乎会不可避免地认为，原始遗传系统经过以基因为特征的进化阶段，在基因达到较长的长度以前，进化时间刚好够编码少数氨基酸，这种观点与基因的外显子理论相一致，它认为真核基因是由外显子和内含子交替出现的不连续结构，反映了基因的进化途径：外显子实际上可能是用于规范原始模块结构的微型基因，通过装配，即加入内含子，从而形成真核基因组中的不连续基因。根据这个理论，真核基因中的每个外显子都对应一个始祖微型基因，而内含子则是组成这些微型基因间的连接点[3]，现在已有大量证据支持这个理论[4]。

始祖基因，即基因外显子理论中假设的微型基因，对它的确认是一个难题，因为自从始祖基因出现以来已经过去了相当长的时间，抹去了它们所有可以探寻的踪迹。但是，概念性的思考和许多基因组的完整序列，使我们现在得以探索这些来自古代进化阶段的古老基因。本章将分析这些基本概念和确认始祖基因时获得的经验性结果，从而明确最后一个全生命共同祖先（LUCA）的性质[1]。

13.2 阐释内含子起源的理论

有三个理论可以用来解释内含子的起源[4~6]。

(1)“内含子后期论”认为，内含子和拼接体首先在真核世系中出现，并从此在真核基因组中积累[2,5]，拼接体是一种细胞器，能从基因转录本中剪切内含子，形成成熟 RNA。新内含子的来源包括“反相剪切”和转座元件的插入[5]。

(2)“内含子早期论”认为，具有内含子-外显子结构的基因在 LUCA 中就存在，甚至可能更早。根据这个理论，原始蛋白质的结构域是随机的，从而促进它们的多样化和进化[3~5]。之后，古菌和细菌域中的内含子丢失了。内含子丢失的分子机制是剪切后 cRNA 的重组；原始反相转录酶可能介导了丢失机制[5]。“内含子早期论”的“内含子侵入”版本，是“内含子早期论”和“内含子后期论”的折中，它认为，内含子出现很早，但在真核生物产生之前并没有广泛分布。它们的重要性仅仅是真核生物进化过程中，大量内含子侵入真核基因的结果。这种侵入刺激了真核生物中的关键事件，如拼接体、细胞核、线性染色体、

端粒酶和泛素信号系统的出现。之后，内含子的进入就被限制了[7]。

(3)“内含子最先论”与“内含子早期论”类似，但回溯的更远，它认为内含子和拼接体是 RNA 世界的遗骸[5,6]。假想的古代 snoRNA 基因常常带有内含子，这个模型与之相符合。由于在蛋白质出现之前，RNA 是唯一能够装配 RNA 原始核糖体的催化剂，因此 snoRNA 应该被用来装配这些原始核糖体[5,6]。因而，根据这个理论，snoRNA 时间上早于编码蛋白质的外显子。RNA 加工在蛋白质出现之前就可能存在，核糖核酸酶 P（RNase P）对 pre-rRNA 和 pre-tRNA 的加工可能就是例子[5,6]。“内含子早期论”和“内含子最先论”都需要已知类型的酶促反应来处理内含子。

13.3　骑行纳古菌的断裂基因和其他特征

骑行纳古菌（*Nanoarchaeum equitans*，Neq）是一种非常特殊的超嗜热古菌，是从一个深海热液口分离出来并得到鉴定的。它是另一种古菌的寄生菌，具有大量十分特殊的性质，如它的 rRNA 并不包含在其他古菌中发现的典型寡聚序列。基于其 rRNA 序列的特殊性，Neq 被认为是古菌的一个新种系，纳古菌门的第一个代表[8]。

Neq 的基因组序列显示，这种生命体具有以下不同寻常的性质。① Neq 具有反常的高数量断裂基因，包括至少 10 个编码蛋白质，6 个编码 tRNA 分子（图 13.1）[9,10]。这 10 个蛋白质编码基因完全断裂成两段，一段负责蛋白质的 N 端，另一端负责 C 端。这些基因占据了基因组中的非连续区域。对 6 个 tRNA 基因也是这样，它们的 5′和 3′部分由完全不同的非连续基因编码。目前，这种编码 tRNA 一半分子的断裂基因只在 Neq 中发现[9,10]，而蛋白质的断裂基因在其他古菌中也发现了，但没有这么高的数量[9]。②Neq 基因组的另一个特殊性质是缺乏操纵子。其他所有古菌都有的操纵子在 Neq 中完全不存在，其相应蛋白质几乎全部存在，但由分散的基因编码[9]。即使是核糖体蛋白质的超操纵子这种最大基因序列，除了一些少数例子外，所有的古菌和细菌都在不同程度上携有，但在 Neq 中却几乎完全缺失。

5′ tRNATrp(CCA)

5′gaatctcgggccggtagctcagcctggttagagcggcggtggccatcccccctttaaattt 3′

3′ tRNATrp(CCA)

5′ggatggccgccgaggctcttaacccgcaggtccggggttcgaatccccgccggcccgtgg 3′

图 13.1　Neq 基因组中 tRNA 半基因的例子。5′端 tRNATrp（CCA）中，框中序列的 1 位与正链 DNA 中的 308、963 号碱基对应，3′端则和反链 DNA 中的 380、364 号碱基对应。

种系发生分析将 Neq 置于 SSU RNA 树上古菌域的最深部[9]，属于广古菌（Euryarchaeota）[11]或泉古菌（Crenarchaeota）的最深分枝[12]。

13.4 有关基因和内含子起源的一种理论

有人认为，DNA 进化较晚，刚好处于 LUCA 的进化阶段之后，即正好处于建立主要系统发生线之时[13,14]。基于这种看法，LUCA 应具有 RNA 基因组。符合这种进化理论的假说之一断言，LUCA 具有由 RNA 片段组成的基因组，而产生基因产物的主要机制之一是反式剪接（trans-splicing）。换句话说，LUCA 的所有 mRNA 都通过反式剪接产生，而物理上连续整体的基因则并不存在[1]。

在这种进化形式下，可以认为 DNA 基因组的形成（在 LUCA 阶段之后出现），是将 LUCA 的片段 RNA 基因组转化为连续的 DNA 基因组，其中每个 RNA 片段对应一个微型基因。这种转化作用将 LUCA 基因组中的所有 RNA 片段都变成小的分散"基因"，即独立的小的关联基因随后会成为同一个基因的片段，分散在 DNA 基因组内。在这个由分散微型基因组成的基因组模型中，每个微型基因对应一个未来真核基因组中的"外显子"，因此，最早的"外显子"在 DNA 基因组中是完全分散的。由此，这个模型需要生命三个领域中不同类型的内含子独立起源，它们的出现极大加速了独立微型基因的结合，形成真核生物中典型的不连续但同链的基因。根据这个模型，在最初的 DNA 基因组阶段，不会有任何内含子参与到基因的装配中来。只是在后来，内含子才在基因装配中起到了基础并且活跃的作用，如同我们今天所见的一样。与假设 LUCA 中存在内含子的基因外显子理论[3,5]相反，这个模型认为内含子出现在 LUCA 阶段之后，或至少没有参与第一个 DNA 基因组的构建，因为这第一个基因实际上是微型基因，编码直接来自 LUCA RNA 片段基因组的模块结构。很明显，如果 Neq 的基因组实际上就是始祖基因的来源，那么就直接支持了基因和内含子起源的这个模型[1,15]。

13.5 关于 tRNA 分子起源和进化的一个模型

有一个构建 tRNA 分子的简单方法（图 13.2）。从一个 RNA 发夹结构开始，直接复制这个发夹，我们就能得到一个十字形结构，它随后的进化就可能产生真正的 tRNA 分子（图 13.2）。tRNA 起源的这个模型得到了广泛讨论[1,15~19]。

支持这个模型的一些最强证据具有历史意义，并与 tRNA 基因中内含子的位置相关。tRNA 起源的这个模型服从基因的外显子理论[3]，它假设 tRNA 基因组中的任意内含子一定位于反密码子环，在那里内含子能够将 tRNA 分子断裂成两半，这与两个单独发夹结构起源相对应。根据基因的外显子理论，这个模型假设，存在编码发夹结构的微型基因，而它们之后的进化则会形成现代 tRNA

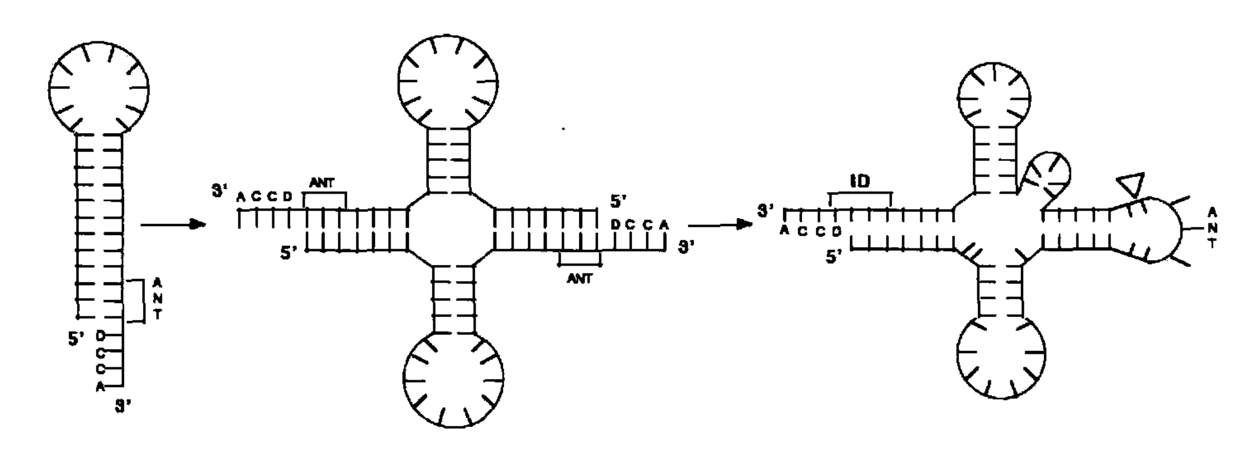


图 13.2　tRNA 分子起源的模型，细节见 Di Giulio 的相关文献[15~17]。ANT 代表反密码子，ID 为同源氨酰 tRNA 合成酶，通过它识别该 tRNA 核苷酸。三角形显示 tRNA 基因中内含子的位置[16,19]。

基因[16~19]。并且，内含子的位置可追溯到 tRNA 基因最初的装配方式上。基于这一点，在所有三个生命领域中，tRNA 基因反密码子环上都有内含子，这是支持这个模型的有力历史证据（图 13.2；见参考文献 1，15～19）。

tRNA 起源的这个模型使它自身得到进一步扩充[18]。已知由微型基因编码的发夹 RNA，其结构使得蛋白质合成的复杂过程能够进行，因而遗传密码的最初组织可能是通过这种简单的 RNA 发夹结构进行的，而不是完整的三叶草 RNA，它可能只是在 LUCA 阶段之后出现的[18]。总之，tRNA 分子的起源很可能不是单一世系的，即编码特定氨基酸的 tRNA 分子可能来源于不同生物种系的不同 RNA 发夹[1,18]。基于这一点，三叶草形 tRNA 基因在进化上应该出现得较晚，可能是在生命三个领域建立过程之中或之后才出现的[1,15,18]。

总之，这种 tRNA 起源理论认为，编码发夹 RNA 的微型基因在进化上出现得很早。

13.6　骑行纳古菌的断裂 tRNA 基因是现代 tRNA 基因进化的始祖

已有证据表明，编码 tRNA 分子 5′和 3′部分的断裂基因是现代 tRNA 基因的始祖（plesiomorphic）形式[1,15]。

（1）tRNA 基因组内含子的 3′端通常位于 tRNA 的 37 号位置，这也是 Neq 的半基因中 tRNA 5′和 3′部分断裂的地方[1,10,15]，这意味着，具有内部插入的 tRNA 半基因和 tRNA 基因一定代表了同一序列的不同进化阶段（图 13.3A；见参考文献 1、15），因为在两个片段上 37 号 3′部分都偶然断裂的可能性极低。如果具有内含子的 tRNA 半基因和 tRNA 基因代表了同一序列的不同进化阶段，那么不带有任何内含子的连续 tRNA 基因一定是这种 tRNA 序列进化的最终阶段。

因此，tRNA 序列进化经历了图 13.3A 中显示的独立阶段，即始祖阶段、过渡阶段和现阶段。这些序列随着时间的推移，使最初始祖阶段的两个分散的断裂 tRNA 基因彼此接近，仅由一个内含子隔开，从而到达过渡阶段，最终像现阶段一样彼此连续，逐渐降低构建三叶草形 tRNA 分子时发生任何错误的可能性[1,15]。相反，基于图 13.3B 中的另一种观点认为，序列进化是从连续的 tRNA 基因，到具有内含子的 tRNA 基因，最后到分散的断裂基因，这将导致错误随进化增加，而不是减少，故这一观点是不合理的。

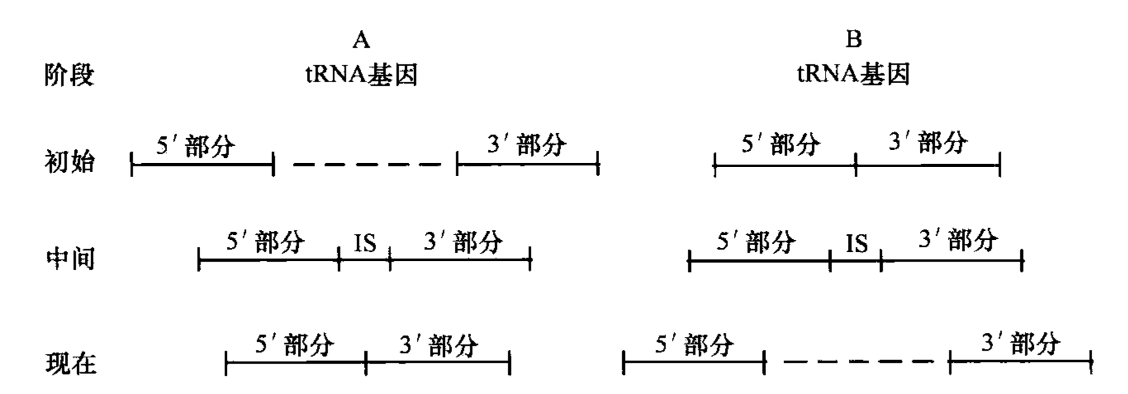


图 13.3 tRNA 基因的两种可能进化序列（IS＝内含子）。A. 进化由断裂基因到非断裂基因进行。B. 进化由非断裂基因到断裂基因进行[1,15]。

（2）对这三个阶段，在大量生命体中发现的实际基因频率也有力地支持了图 13.3A 中的进化顺序，而不是图 13.3B。具有连续 tRNA 基因的物种远远多于带有内含子的 tRNA，更远远多于具有断裂 tRNA 基因的生命体，这与图 13.3A 中分配给断裂 tRNA 基因的残缺状态完全一致。反之，生命世界中的多数 tRNA 基因都是进化后的状态。残缺状态非常少见，因而否定了连续 tRNA 基因是始祖的看法，支持断裂的 tRNA 基因是始祖的观点[1,15]。

总之，Neq 中编码 tRNA 分子 5′和 3′部分的半基因代表了 tRNA 基因的始祖形式，从中可能演化出现代 tRNA 基因[1,15]，它们同时也是基因外显子理论所预测的微型基因。

13.7 tRNA 基因多种系起源法则

关于 Neq 断裂 tRNA 基因的观点，可以归纳为 tRNA 基因多种系起源法则："如果 Neq 的 tRNA 半基因代表了 tRNA 基因的始祖形式，那么在生命体中观察到这种基因的存在，意味着 tRNA 的单种系起源假说是错误的，而多种系起源是正确的。"

要证明这个法则的证据十分简单。如果单种系起源假说正确，那么从 LUCA 的进化阶段开始，我们就应该能够明显观察到，只有完整的 tRNA 基因，而没有半基因，这是因为在 LUCA 阶段之后的任何时间，单种系起源都不允许同

时存在连续和断裂这两类 tRNA 基因，由此可见，在 LUCA 阶段之后对 tRNA 半基因的发现，和这些半基因是始祖形式的证据，都意味着单种系起源假说是错误的[1,15]。由此可见，tRNA 基因是多种系起源，LUCA 一定含有 tRNA 半基因。

13.8 一些骑行纳古菌蛋白质的断裂基因可能是始祖基因

如果 Neq 中一些 tRNA 的 5′和 3′断裂半基因和它的 10 种不同蛋白质断裂基因没有关系，那将是很令人吃惊的[1,9]。我们的观点是，编码一些 Neq 蛋白质的断裂基因也是这些基因的近祖形式。接下来的观察结果能有力支持这个观点。若将由 Neq 中不连续断裂基因编码的两个蛋白质片段与真核生物中的同源连续蛋白质对比，我们应该能够预测这些真核基因中内含子的位置。换句话说，如果如基因的内含子理论预测的那样，内含子在基因装配中起重要作用，那么在多重比对区域中，真核基因内部，Neq 的 N 端片段末端和 C 端片段起始处就应该有内含子。实际上，对所有 6 种蛋白质基因都能检验这个预测，预测位点上具有内含子，从而证实了假说（手稿正在准备）。因此，这个观察结果使我们更加相信，在 LUCA 进化阶段只有完全断裂的基因，它们后来结合，正如检验到的那样，在连接位点具有内含子。因此，我们只能得出的结论是，编码一些 Neq 蛋白质的不连续断裂基因是这些基因的近祖形式，与断裂的 Neq tRNA 半基因一致。

13.9 蛋白质基因的多种系起源

tRNA 基因的多种系起源法则显示了在 LUCA 阶段，存在完全分离的 tRNA 基因，编码分子的 5′和 3′部分[1,15]。这与对 Neq 中一些编码蛋白质的断裂基因的观察结果一起，指出了一个更普遍的状态，即所有基因都分裂成片段，必须装配成 mRNA。这种普遍状态可能应用于 LUCA，因而现存生命的所有连续基因都应该是在生物多样性的主线建立之后产生的，意味着它们的起源也是非单一种系的[1,18]。总之，LUCA 的 mRNA 可能是通过反式剪接形成的，从而翻译成蛋白质产物，它的基因实际上进化得更晚，即刚好在生命领域建立之后出现。有大量证据来证实基因这种非单种系的起源方式[1,18]。关于这点，值得一提的是在生命的三个领域中，复制途径似乎不是同源的[13]，这与其他观察结果一起，显示 DNA 刚好在主要的种系建立之后出现，即在 LUCA 阶段之后出现[13,14]。这与基因的多种系起源假说完全契合：DNA 的较晚起源可能与多种系一致，也就是基因起源较晚。DNA 出现在 LUCA 阶段之后，在每个主要的种系中都出现，从而在这些种系内完成独立的、“确定的”起源，因此，我们似乎必须慎重考虑所有基因多种系起源的可能性。

13.10 结论和展望

很明显，骑行纳古菌的断裂基因，尤其是编码其 tRNA 分子 5′和 3′部分的那些基因，是我们已知基因的最古老形式，这意味着生命体保留着有关生命早期进化阶段的秘密，即 tRNA 起源的最早阶段始祖基因的特性，如断裂 tRNA 基因代表着微型基因[1,15]。这使我们有可能精确重建有如 tRNA 起源这样的原始事件，以及对其发生时间进行确定。因此，打开迷人的前景，随着新理论的发展和新基因组的测序，将能探索那些长期作为科学问题存在的古老事件，如蛋白质合成的起源。更通俗地说，我们有信心揭示生命进化的早期阶段，而通过对生命体的研究，将能发现生命起源的迷人答案。

参考文献

扩展阅读

1. Di Giulio M. The nonmonophyletic origin of the tRNA molecule and the origin of genes only after the evolutionary stage of the last universal common ancestor (LUCA). J Theor Biol 2006; 240:343-352.
2. De Souza SJ. The emergence of a synthetic theory of intron evolution. Genetica 2003; 118:117-121.
3. Jeffares DC, Mourier T, Penny D. The biology of intron gain and loss. Trends Genet 2006; 22:16-22.

专题文献

4. Gilbert W. Why genes in pieces? Nature 1978; 271:501.
5. Roy SW. Recent evidence for the exon theory of genes. Genetica 2003; 118:251-266.
6. Poole AM, Jeffares DC, Penny D. The path from the RNA world. J Mol Evol 1998; 46:1-17.
7. Koonin EV. The origin of introns and their role in eukaryogenesis: a compromise solution the introns-early versus introns-late debate? Biol Direct 2006; 1:22 doi: 10.1186/1745_6150-1-22.
8. Huber H, Hohn MJ, Rachel R et al. A new phylum of Archaea represented by nanosized hyperthermophilic symbiont. Nature 2002; 417:63-67.
9. Waters E, Hohn MJ, Ahel I et al. The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism. Proc Natl Acad Sci USA 2003; 100:12984-12988.
10. Randau L, Munch R, Hohn M et al. Nanoarchaeum equitans creates functional tRNAs from separate genes for their 5′- and 3′-halves. Nature 2005; 433:537-541.
11. Brochier C, Gribaldo S, Zivanovic Y et al. Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? Genome Biol 2005; 6, R42 doi:10.1186/gb-2005-6-5-r42.
12. Wong JT, Chen J, Mat WK et al. Polyphasic evidence delineating the root of life and roots of biological domains. Gene 2007; 403:39-52.
13. Forterre P. Three RNA cells for ribosomal lineages and three DNA virus to replicate their genomes: A hypothesis for the origin of cellular domain. Proc Natl Acad Sci USA 2006; 103:3669-3674.
14. Poole AM, Logan DT. Modern mRNA proofreading and repair: clues that the last universal ancestor possessed an RNA genoma? Mol Biol Evol 2005; 22:1444-1455.
15. Di Giulio M. Nanoarchaeum equitans is a living fossil. J Theor Biol 2006; 242:257-260.
16. Di Giulio M. On the origin of the transfer RNA molecule. J Theor Biol 1992; 159:199-214.
17. Di Giulio M. Was it an ancient gene codifyng for a hairpin RNA that, by means of direct duplication, gave rise to the primitive tRNA molecule? J Theor Biol 1995; 177:95-101.
18. Di Giulio M. The nonmonophyletic origin of tRNA molecule. J Theor Biol 1999; 197:403-414.
19. Di Giulio M. The origin of the tRNA molecule: implications for the origin of protein synthesis. J Theor Biol 2004; 226:89-93.

14 遗传密码

王子晖

14.1 引言

生物体含有非常大的信息量。信息主要由 DNA 中的多聚序列携带，可以通过转录和翻译被转化为 RNA 和蛋白质。这些多聚物由多种单体结构单元组成，信息就体现为不同单体的序列排列方式及其长度。类似地，异源多聚物信息系统也以语言中字母和文字序列或音乐中听觉符号序列的方式存在。一个序列中的最大信息量为 I_{max}，以比特形式表示为

$$I_{max}=N\ \mathrm{Log}_2\ (M)$$

式中，M 为单体类型数；N 为序列长度。复杂化合物如叶绿素和原血红素将信息储存在其原子和键的空间分布中，但这些化合物难以复制，并为信息量的上限所限制。比较而言，多聚物信息量实际上可以通过延长序列来达到无限。人类基因组有 4 个单体和 30 亿碱基对长度，其信息量为

$$I_{max}=3\times10^9\ \mathrm{Log}_2\ (4)\ =6\times10^9\ \text{比特}$$

这样的信息量相当于 400 册的人体建筑手册，每册 600 页，每页 5000 个 5 比特的字母。这还远未接近 DNA 的上限。肺鱼基因组包含的 DNA 就远多于人类。

由于 DNA（T，C，A，G）和 RNA（U，C，A，G）都使用 4 单元表，DNA 单元可以与 RNA 单元一一对应。翻译时，正如 George Gamow 所预见的[6]，每个氨基酸和相应核苷酸间需要有一定联系。基于这一点，遗传密码可以将 64 个三联密码子分配给氨基酸和终止信号，将 RNA 语言翻译成蛋白质语言。在电报传递中，莫尔斯码将字母表中的字母编码为“点”和“线”。这完全是一种随意编码，很容易被其他密码取代；考虑到安全问题，军方传递放弃了莫尔斯码而改用其他密码。由于遗传密码也是随意的，它在所有生命体中的广泛利用意味着在距今约 36 亿年，现今生命体的最早分支出现之前，这套密码就已经确立了（图 14.1）。它就像最早的岩石一样古老，也可能更老。

地球形成	稳定的水圈	前生物化学	前RNA世界	RNA世界	遗传密码 第一个DNA/蛋白质生命	LUCA生命的多样化
4.5	4.2	4.2-4.0	~4.0	~3.8	~3.6	3.6-至今

图 14.1　上亿年来早期生命的历史时间轴。根据 Joyce 的相关文献[3]。

在埃及 Rosseta 港口发现的 Rosseta 石创作于公元前 196 年。其上用埃及象形文字、埃及通俗语和希腊语雕刻着同一段话。这段文字为托勒密 5 世关于来自葡萄园的葡萄酒及其他重要政府大事的法令，对它的翻译使得 Jean-Francois Champollion 得到解读象形文字的规则，并揭开古埃及历史的面纱，极少有政府法令能够得到这么永恒的用途。同样，遗传密码刻出了两种语言，即核酸和蛋白质语言（图 14.2）。可以预期，对遗传密码结构的解释将为蛋白质语言起源带来无可取代的理解，目前尚未有任何生化考古领域的研究，可像遗传密码一样，得到几十年来以多种手段进行的持久分析。

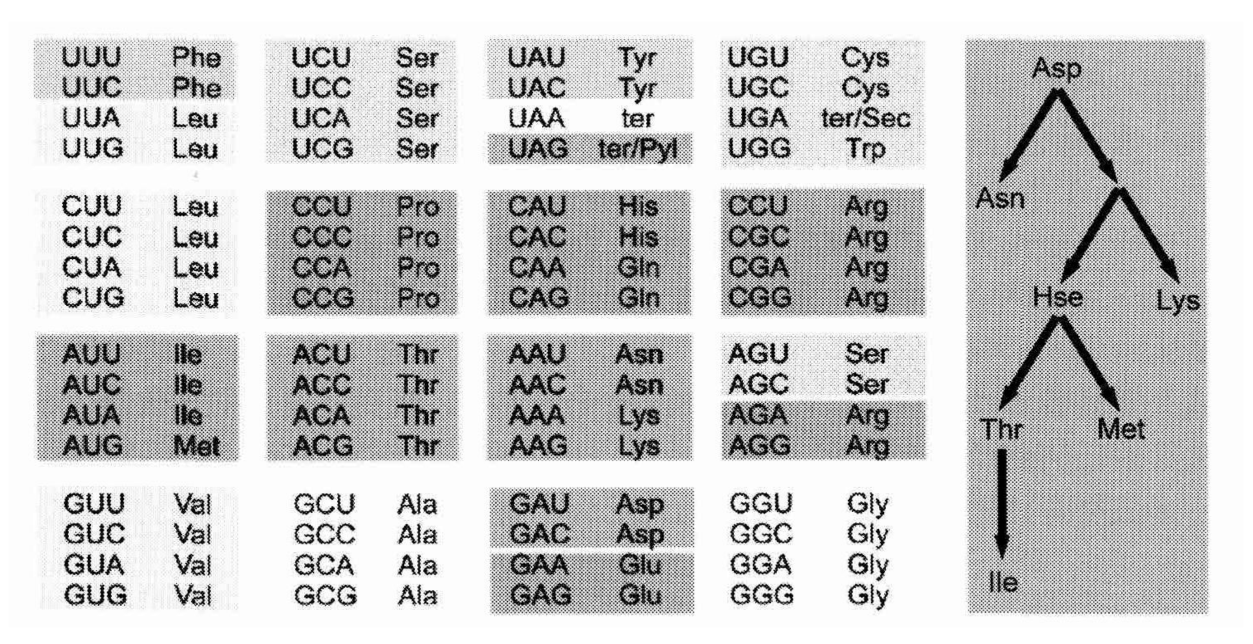


图 14.2 全生命体系遗传密码。来自不同生物合成途径的氨基酸家族密码子用不同彩色标出。吡咯赖氨酸（Pyl）只部分使用 UAG，硒代半胱氨酸（Sec）部分使用 UGA。天冬氨酸合成家族在右侧给出（见彩插）。

14.2 密码的诞生

14.2.1 对侧链的需求

核酶、RNA 核酸适体、酶和结合蛋白都依赖于一个最优序列来进行三维折叠，使关键侧链最优化地布置于底物或配体周围，因此，要获得一个有效侧链的组合就变成进化的关键需求。在 RNA 世界中[7]，RNA 分子需要行使信息复制子和核酶/核酸适体双重功能。通过互补配对复制时，两对互补的嘌呤-嘧啶碱基很是理想，仅有一对碱基信息量不足，三个或三个以上的碱基配对又可能通过错读增加复制错误。可是，两类碱基只会给 RNA 带来 4 种结构单元，这严重限制了核酶侧链的灵活性及核酸适体活性。对于这一复制和功能间的矛盾，一种解决方法是限制碱基数目，使复制错误保持低水平，但是同时引入转录后修饰，增加侧

链多样性[8]。由于修饰过的核苷酸不能参与复制，它们也不会带来复制错误。目前，在 tRNA，rRNA，snRNA 和 mRNA 及生命的所有三个领域中，发现了 95 类修饰[9]。除了 3′-CCA 端，tRNA 上达 80％的核酸位置都被修饰[10]。而且，基于古代 8 个物种得到的全生命共同祖先（LUCA）基因组亦包含这样的 RNA 修饰酶，如二甲基腺苷酸转移酶、假尿苷合成酶、喹啉/archaeosine tRNA 核糖基转移酶、硫代腺嘌呤合成酶和 tRNA 1-甲基腺苷甲基转移酶，证明了 RNA 修饰的极早期起源（附录 15.1 和参考文献 11）。但是，每种修饰都需要修饰核酶，它们必须能够将其底物核苷酸与其他核苷酸区分开来[12]。随着这些 RNA 修饰系统变得越来越累赘，进化出替代的生物多聚物作为催化剂和底物结合物的需要也随之增加。因为这些替代物不具有信息遗传的作用，所以它们侧链多样性的增加不会影响遗传信息的稳定。在选择多肽前，RNA 世界很可能试用了不同的替代物。最后，RNA 指导多肽合成，并使其具有催化和结合配体的重要能力。

选择多肽作为替代多聚物并不奇怪。前生物环境中的氨基酸侧链具有无与伦比的催化和配体结合能力。蛋白质结合配体具有超凡能力，例如，IgG 抗体上的可变区由 216 个氨基酸残基构成，一半是 L 链，一半是 H 链，它们能够产生多种具有明显特异性的抗体。而在目前的催化研究中，蛋白质突出的催化能力是不能被其他任何物质所超越的。核酶或酶（C）的催化作用是通过催化剂和底物（S）形成复合体进行的，之后是底物向反应产物的转化和游离催化剂的再生。

$$C+S \rightleftharpoons CS \longrightarrow C+\text{反应产物}$$

催化反应的速率 v 由 Michaelis-Menten 方程描述：

$$v=k_{cat}\cdot[C_0][S]/([S]+K_m)$$

催化速率常数 k_{cat} 决定了 S 在总催化剂浓度 $[C_0]$ 时转化成产物的最大速率，而 Michaelis 常数 K_m 决定了催化剂结合位点半饱和时所需的 S 浓度。RNA 中许许多多的氢键使得它能够折叠出各种构象，为所有配体提供高亲和位点，使核酶 K_m 通常较低。但是，由于只有 4 种侧链，它们一般难达到较高的 k_{cat}。比如，RNase A、RNase T1 和 RNase T2 这几种水解酶的 k_{cat} 为 5000～180 000/min，而剪切 RNA 的核酶，包括 1 类核酶和锤头状核酶，其 k_{cat} 通常为 100/min 或更小[13~15]。RNA 世界可能持续了不到 10 亿年（图 14.1），而 DNA-蛋白质世界已经超过了 30 亿年，并仍在发展壮大。因此，与 DNA-蛋白质世界相比，RNA 世界肯定有一些潜在的缺陷，不能达到较高的 k_{cat} 可能是其中关键的一点。

14.2.2 RNA 多肽化

尽管由核酶向肽酶转变的进化优势很明显，但这种转变的实现也需要满足两个要求。第一，蛋白质的形成必须严格服从 RNA 的指令，这样 RNA 在进化中

积累起来的有用信息才不会被浪费。系统难以承受两次基因信息的重复发展。之后，还需要发展出一套辞典或密码，将 RNA 语言翻译为蛋白质语言，有多条途径可以实现这个过程。编码单元（密码子）可能包含重叠或不重叠的核苷酸，每个密码子亦可能包含一个、两个、三个或更多的核苷酸。就算不是全部，大多数这类途径都很可能被核糖生命体尝试过，剔除了那些不合格的途径后，最终留下非重复三联密码子作为最优选择，保持了足够的氨基酸多样性但又不会过于累赘。第二，在脊椎动物阶段之前，未雨绸缪这个性能还没有在生物系统中出现，因此，遗传密码发展的每个阶段都需要有眼前的好处推动着。一个符合这两个要求的发展机制为二阶段的 RNA 多肽化机制[16]（图 14.3）。

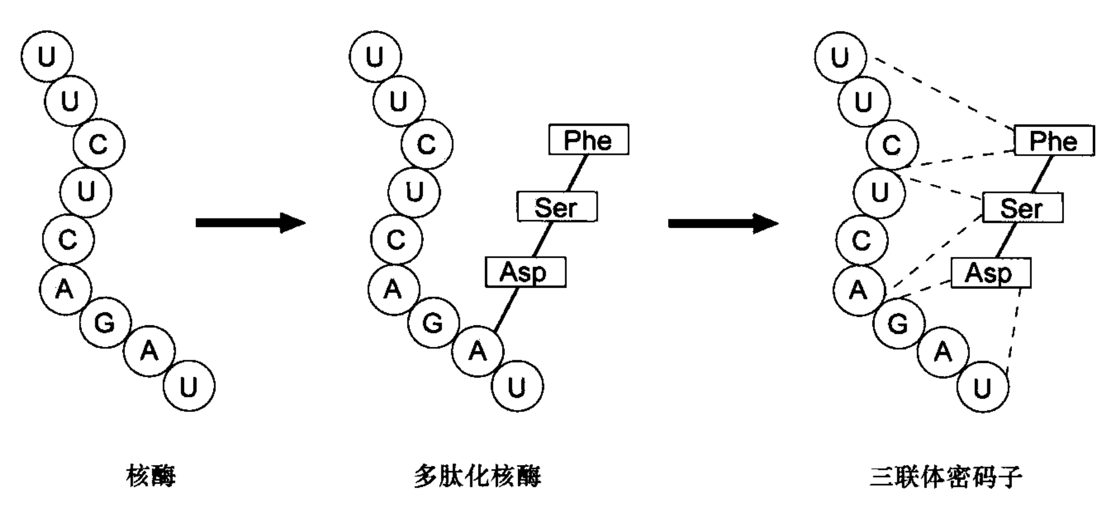


图 14.3 三联密码子通过 RNA 多肽化机制的发展。

RNA 多肽化机制的出发点（图 14.3 左侧）是一个没有任何复制后修饰的核酶片段。之后，各种引入的后修饰，向核酶中添加了其他的侧链。由于前生物环境中有可以利用的氨基酸，它们就有可能被包含在添加的修饰当中。RNA 中氨酰基和肽基侧链的结合可以得到如下支持：现今的有机体利用了不同的含有多肽的核苷酸分子（表 14.1），RNA 核酶对 RNA 的氨酰化作用[17~19]，以及氨基酸和多肽对核酶的激活[20,21]。核酶性质的非核糖体多肽合成酶（NRPS）系统，例如，现今可以合成多肽的多肽合成酶[22]，可以催化 RNA 上多肽键的形成。RNA 核苷酸序列指导了其上形成多肽的氨基酸序列，使不同的氨基酸匹配同源核苷酸，形成第一阶段的肽辅核酶（图 14.3 中间）。在黄素蛋白中，黄素辅酶通常与酶紧密结合，如 D 型氨基酸氧化酶；亦可能与酶可逆分离。同样，在有些转录后修饰中，多肽和氨基酸也能够与 RNA 共价结合，或可逆结合[23]。通过多肽与核酸的共价结合，可以避免多肽因子与其他核酶意外结合，而在催化位点周围构建一群多样的侧链，就像今天在大多数酶的激活位点那样。

表 14.1　含与核苷酸或有机碱相连的多肽或氨基酸的化合物

化合物	碱基-氨基酸连接	功能
甘氨酰 tRNA	O-肽核苷	细菌细胞壁合成
UDP-N 乙酰胞壁酰五肽	UDP-GNA-乳酸基肽	细菌细胞壁合成
辅酶 A	ADP-泛酸-半胱胺	酰基转移
叶酸	蝶酰基多聚谷氨酸盐	转移一个碳原子
420 因子	二葡萄糖类黄酮	转移一个碳原子
二氧化碳还原因子	吩噁嗪酮-双环五肽	转移一个碳原子
放射菌素	核糖酸酪氨酸	抗生素
腺苷酸蛋白	双环五肽	调节蛋白
核苷核糖化蛋白	N-C-配糖	调节蛋白
病毒 RNA 蛋白	骨髓灰质炎病毒 $\alpha_0$5′-蛋白-pUUAAAACAB	RNA 合成前体
N-肽-tRNA	嘌呤-6-甲氨酰-苏氨酰胺基-基因	转运 RNA
O-肽-tRNA	O-肽核苷	介导蛋白质合成

即使是短多肽链，如 14 个氨基酸长度，也可具有催化活性，而在随机多肽序列当中有活力的比在随机 RNA 序列中找到的多 10^6 倍[8]，因此，RNA 上多肽侧链的催化功能很快并毫无疑问地证明了自己。于是它们与 RNA 分离，开始行使原始的酶的功能，取代了核酶。这时，进化进入了三联编码阶段，指导氨基酸合成的同源三核苷酸 RNA 片段成为三联密码子（图 14.3 右侧）。不同的适用 RNA 会被不同的氨基酸氨酰化，成为现代氨酰 tRNA 家族的前体[3]。这些可被氨酰化的 RNA 可以成为单茎环微螺旋[24,25]或精巧的氨酰化核酶[3]。再之后，DNA 取代 RNA 成为基因，RNA 分子被限制于现今与蛋白质合成相关的角色，如 rRNA、tRNA、mRNA、催化核糖体的多肽键形成，以及为结构 RNA 加工的核酶。

14.2.3　密码的扩充

一旦遗传密码建立，驱动多肽侧链多样化的需求就开始在遗传密码层面促进氨基酸多样性的增加，使其不受前生物环境中只能利用 1 期氨基酸的限制。这个方向的唯一可行方式就是在细胞中发展出新的生物合成途径，为遗传密码产生 2 期氨基酸[26~28]（表 1.1）。这些 2 期氨基酸向蛋白质侧链中加入了苯基、吲哚、咪唑、巯基、酰胺和正离子基团。咪唑很有催化效力，比如，一个简单的 Ser-His 二肽就能够催化 DNA、蛋白质和酯的裂解[29]。

如专题 14.1 所述，所有的字母语言都需要一套合适的字母去表达人类声音。一旦得到了合适的一套字母，字母表就确定了下来。同样，在试用了新的 2 期氨基酸，并选择了那些最能够满足需要的氨基酸之后，遗传密码在未来的 30 亿年

内就被确定了。确定前的 2 期氨基酸字母的增加，事实上为蛋白质在未来世代的功能表现建立了卓越的基础。在遗传密码扩充过程中，一个新氨基酸的加入，引入了新的品种，同时也为蛋白质顺序带来了干扰。例如，通过翻译前（pretran）合成引入谷氨酰胺（图 1.4）在 tRNA 上取代谷氨酸，承接了 CAA-CAG 密码子，密码也就获得了新的氨基酸。过程中整个基因组内所有 CAA-CAG 编码的谷氨酸残基都替换为谷氨酰胺，这显然是一种干扰。正如人们在图书馆中会考虑要不要大声说话，但在市场上就不会如此一样，噪声在高噪声背景中比较容易被接受。同样，当背景中翻译错误率高的时候，新氨基酸带来的干扰对进化的不利影响较小。这样，当翻译错误率低的时候，添加新氨基酸，扩充密码子的过程也就由于干扰相对大而随之停止了[28]。因为翻译错误率是基于翻译机制中蛋白质的功能表现的，所以只有当所有这些蛋白质都表现出色的时候，才能有足够低的错误率使密码子确定下来。因此，除非全套氨基酸编码达到完美，不然密码子的进化就永远不会停止。由此可见，蛋白质分子的完美功能并不是意外获得或幸运造成的，而是密码扩充过程中错误反馈严格控制的结果。

专题 14.1 字母表的演化

在人类语言中，字母表中的字母表达了人类的不同声音。由于人类能发出 40 种不同的基本声音，一个有效的字母表不需要超过 40 个字母。由于有些时候字母的组合能代替新字母，表达一种发音，如“sh”，“ch”，“oe”等，因此，不同字母表的字母数可以不同。

斯拉夫语	33
英语	26
希伯来语	22
古匈牙利语	39
国际音标	40
拉丁语	23

字母表演化的两种特性与蛋白质字母表尤其相关。首先，字母的数量必须足够，因为太少字母不足以表达人类发音，会更像咕哝声而不是语言。其次，一旦字母表建立，就趋于恒久不变。字母表是文化遗产，任何改变其字母的尝试，如让英语字母表的字母数与斯拉夫语相同，或反过来，都会导致相当大的阻力。

同样，蛋白质字母表中的氨基酸也必须达到一定的数目和平衡，从而支持蛋白质的高效功能。比如，根据方程 $I_{max}=N\log_2(M)$，通过使 N 达到很大，仅仅由 Gly、Ala 和 Val 这三个氨基酸组成的简化蛋白质也能产生大量信息，不过这些信息不会很有用。无论蛋白质多长或其序列如何重排，这种简化蛋白质能催化的反应仍极其有限，因为它缺乏了十分重要的侧链多样性。由于蛋白质字母表是蛋白质组中遗传信息的基础，也是细胞遗产，同样也趋于恒久不变。改变目前这个字母表将向蛋白质序列中引入大量无意义的干扰，因而会产生很大阻力。

在前生物及早期生物进化中，侧链需求是最有力并永不停息的驱动力。从

RNA 世界开始，这些需求先后引入了转录后修饰、肽基核酶、1 期遗传密码、2 期遗传密码，并在半确定的遗传密码中最后引入了硒代半胱氨酸（Sec）和吡咯赖氨酸（Pyl）（图 14.2），甚至当密码子完全确定之后，翻译后修饰扩增（PTM）还向蛋白质中引入了更大量的氨基酸侧链（表 14.2）。单单 PTM 糖基化就包括了糖类与丝氨酸、苏氨酸的 *O* 型连接，或与天冬酰胺的 *N* 型连接，线性及支链的形式，并包括了甘露糖、半乳糖、岩藻糖、乙酰氨基半乳糖、乙酰氨基葡萄糖和唾液酸。

表 14.2　一些翻译后修饰

乙酰化作用	ADP-核糖基化作用
甲基化作用	黄素连接作用
C 端酰胺化作用	氧化作用
生物素化作用	棕榈酰化作用
甲酰化作用	磷脂酰肌醇连接作用
γ-羧化作用	磷酸泛酰巯基乙胺化作用
谷氨酰基化作用	磷酸化作用
糖基化作用	焦谷氨酸合成
糖基化作用	辅氨酸外消旋作用
血红素连接作用	精氨酸化作用
羟基化作用	硫化作用
碘化作用	ISG15 蛋白连接作用
异戊二烯化作用	SUMO 蛋白连接作用
异戊烯化作用	泛素化作用
豆蔻酰化作用	精氨酸转化成瓜氨酸
法尼基化作用	二硫键
香叶酰香叶酰化作用	蛋白质水解裂解作用

14.2.4　密码进化

遗传密码是合理性与不合理性的奇妙混合，它的合理性表现在其有序的分配，主要由 4 个密码子对应一个氨基酸或两个氨基酸，而其不合理的地方是比如丝氨酸 6 个密码子处于两处不同位置。密码的这种奇特结构及其全生物界通用性曾被看做是“意外冻结”（frozen accident）的结果[31]，但是目前至少三种改进密码的进化机制已经被认识到了。

错误最小化

密码的一个特征是物理上相似的两个氨基酸，其密码子也常常在编码上相邻，相邻的意思是，任意两个密码子间只有一个碱基不同。例如，亮氨酸（Leu）的CUU-CUC-CUA密码子，与异亮氨酸（Ile）的AUU-AUC-AUA密码子和缬氨酸（Val）的GUU-GUC-GUA密码子相邻，而这些氨基酸都具有庞大的疏水侧链。其结果是，若将CUU错读为AUU或GUU，对蛋白质结构只会有微小影响，这样的密码子分配方式能够减少错误带来的影响，在进化中占有优势，尤其是在原初阶段，复制、转录和翻译错误频繁发生，这种错误最小化的模式不仅在Leu-Ile-Val中存在，同时也在其他氨基酸中存在，如Ser-Thr、Phe-Tyr、Asp-Glu和Lys-Arg[32~36]。

但是，密码中有三种非优化（Non-Op）类型Ⅰ～Ⅲ，其对错误最小化原则的违背是不言自明的。

（1）Non-Op Ⅰ：任意两个氨基酸物化性质的不同都可以用Grantham“化学距离”方程，结合组分、极性和体积这三个参数计算出来。在20个编码氨基酸形成的190对组合中，Cys-Trp之间的化学距离为215，是Leu-Ile间化学距离5的许多倍[37]。尽管如此，Cys的UGU-UGC密码子仍与Trp的UGG密码子共享UGN盒。

（2）Non-Op Ⅱ：错误最小化的极限就是当一个密码子被错读为其同义密码子时发生的零错误，因此，错误最小化的第一步最大可能就是在相邻位置上尽量放置同义密码子。密码中大多数密码子都是这样。但是，Ser的AGU-AGC密码子却不与任何Ser UCN密码子相邻，这显然与错误最小化背道而驰。

（3）Non-Op Ⅲ：在需氧的大肠杆菌K12中，Met密码子的总利用频率是2.48%，Cys的是1.39%，而在厌氧的*Methanopyrus kandleri*（与LUCA相近的物种：见15.2节）中，Met是1.66%，Cys是1.31%，但是Met仅被分配到了一个密码子而Cys有两个，这样不合适的密码子分配根据任何物理准则都意味着密码的非最优化。

现在，由共进化理论得知，Non-Ops Ⅰ～Ⅲ并不意味着放弃或背离错误最小化。UGN盒曾经属于Ser，与它的UCN和AGY密码子形成邻接域。但是，当Ser通过生物合成产生Cys和Trp之后，它将UGN密码子放弃，让给了Cys、Trp和终止信号。再之后，Sec——另一种Ser衍生氨基酸，亦成功得到了UGA，从而与同样是Ser衍生物的Cys和Trp进入了同一个密码盒。Ser将UGN放弃给其合成产物的行为扰乱了Ser密码子域的邻接性，同时造成了Non-Op Ⅰ和Non-Op Ⅱ。Trp和Met只得到了一个密码子，这意味着这两个氨基酸在密码扩充过程中来得较晚。因此，Ser对UGN密码子的放弃显然不会在LU-

CA 出现及密码子冻结之前太早发生，为进化留下足够的适应时间去修复 Non-OpⅠ和 Non-OpⅡ以减少错误。同样，Met 和 Trp 的较晚到来，使其不会有足够的时间去好好调整它们分得的密码子数量，从而引起 Non-Op Ⅲ。出于同样的原因，对于密码中其他出现较晚的变化，也不会有足够时间在冻结之前优化调整以减少错误。

除了调整时间不够外，还有一个事实限制了错误最小化，即每个密码子都有 6 个邻居，即使只算前两个碱基，每个碱基位置都有三个邻居。这样对六个邻居中一部分进行适应使错误减少，往往就带来对其他邻居适应的恶化。概括地说，密码错误最小化达到的程度为 40％～45％[38,39]，它为全生命通用编码的进化贡献了 10^{-6} 的选择幅度[34]。

立体化学相互作用

当一个氨酰基 tRNA 化合物定位于核糖体上，用于多肽键形成时，与其 3′端相接的氨基酸有可能与密码子及反密码子相距太远，不易与它们进行直接的物理相互作用。但是，对于原初 tRNA 微螺旋及当密码子或反密码子位于 tRNA 接纳茎（acceptor stem）上的时候，直接互作会较容易实现[40]。而且，氨基酸及其 tRNA 上的反密码子可能同时与 aaRS 的活性位点结合，因此，它们的相互作用可能直接在 aaRS 上进行或间接通过 aaRS 进行。例如，一个疏水的 aaRS 活性位点可能会优先结合一个疏水氨基酸及一个具有疏水反密码子的 tRNA，从而促进了疏水氨基酸和疏水反密码子间的联系[5]。因此，直接和间接的立体化学相互作用都会引致疏水氨基酸和它们的密码子/反密码子间建立实验可观察到的联系[32,33,41,42]。

RNA 核酸适体能够结合 Trp、Arg、Val、Ile、Tyr、Phe 和 His、Trp 或 Leu 的氨基酸结合区中已被发现含有其所结合氨基酸的同源三联密码子或反密码子[43,44]，这个发现支持了疏水联系的显示，即氨基酸-密码子/反密码子相互作用显著影响了密码子分配[41~47]。尤其在起初密码子被分配给 1 期氨基酸期间，翻译错误和氨基酸共进化都还没有成为重要指导因素之前，立体化学相互作用机制就可以为全生命通用编码的进化贡献了 0.04％或 4×10^{-4} 的选择幅度。

氨基酸生物合成

由不同氨基酸生物合成家族在密码里所占密码子的位置，可以看到氨基酸生物合成和密码子分配之间的相关性[48,49]。该相关性建立的机制，将在下一节讨论。

14.3　共进化理论

共进化理论（coevolution theory，CET）描述了一种机制，即通过生物合成

产生的 2 期氨基酸进入密码系统，以及密码子从环境中得到的 1 期氨基酸转移给 2 期氨基酸的机制："密码子系统的结构基本是氨基酸前生物形成途径的印记，该印记目前仍可以在氨基酸生物合成途径识别出来。因此，遗传密码的进化可以用氨基酸生物合成中前体-产物的关系来阐明。大多数氨基酸前体-产物对的密码子区域都应该是相邻，被单个碱基的最低差异分隔开的[26]。"

CET 的证据的来源很广[5,50~52]，可以通过 4 个独立的并经得起测验的原理来说明。

原理 1：在生命初始时，前生物环境不能提供所有 20 种标准氨基酸，但可以通过创新性的生物合成（见 1.4.3 节）进行补充。

若能在合适的前生物条件下合成 20 种标准氨基酸，或者能在陨石上发现它们的全部，得到它们地外合成的证据，那么这条原理就将被推翻。但自 CET 提出 30 年来，这两方面都找不到不利于原理 1 的证据。相反，遗传密码结构、大气氨基酸合成和陨石氨基酸这三项证据的吻合，无疑地显示出前生物只能利用环境中产生的 1 期氨基酸，而不能利用生物合成的 2 期氨基酸（表 1.1）。

1 期：Gly、Ala、Ser、Asp、Glu、Val、Leu、Ile、Pro 和 Thr

2 期：Phe、Tyr、Arg、His、Trp、Asn、Gln、Lys、Cys 和 Met

此外，根据蛋白成分年表推论出的氨基酸进入蛋白质的先后次序，9 个最古老的氨基酸——Gly、Ala、Val、Asp、Pro、Ser、Glu、Leu 和 Thr 均为 1 期氨基酸[53]，这更进一步支持 1 期氨基酸进入密码是在 2 期氨基酸之前。

由于 2 期氨基酸 Gln 和 Asn 非常不稳定，它们在前生物环境中的浓度分别不可能超过 3.7 pmol/L 和 24 nmol/L 的微量，所以在前生命环境中去寻找它们是没有意义的（见 9.2 节）[50]，这已足够证明原理 1。

原理 2：翻译前（pretran）合成为一些 2 期氨基酸的编码提供了机制。

在不同有机体中，Gln、Asn 和 Cys 进入蛋白质的过程，是由两条途径介导的：要么通过 GlnRS、AsnRS 或 CysRS，要么通过 pretran 合成，由 Glu-tRNA 生产 Gln-tRNA（图 1.4），Asp-tRNA 生产 Asn-tRNA，或由磷酸 Ser-tRNA 生产 Cys-tRNA。若能证明 aaRS 的使用早于这几个例子的 pretran 合成，则 pretran 途径仅出现于进化中较晚阶段，与前 LUCA 时代无关，那么原理 2 就被推翻。但是，多条线索都指出，LUCA 酷似甲烷火菌（*Methanopyrus*）。由于甲烷火菌中没有 GlnRS、AsnRS 和 CysRS[5,11,14]（附录 15.1），也就是说在 LUCA 阶段就有了翻译前合成，发生于 GlnRS、AsnRS 和 CysRS 之前。此外，比较系统发生学亦指出，翻译前合成出现在 GlnRS 和 AsnRS 之前，与 CysRS 同时[55,56]。Sec 进入蛋白质同样也依赖翻译前合成[57,58]。这些发现证明，翻译前合成的使用在 aaRS 之前，至少对 Gln、Asn、Sec，很有可能对 Cys 也是这样。这证明了原

理 2。

对原理 2 的证明同时也是原理 1 的又一证据。既然 Gln、Asn 和 Sec 需要通过翻译前合成进入蛋白质，显然它们并不能从环境中获得。

原理 3：氨基酸之间的生物合成关系是决定密码子分配的重要因素。

氨基酸生物合成和密码子分配间的显著联系包括：①Ile、Met、Thr、Asn 和 Lys，同属于 Asp 生物合成家族，其密码子位置都位于密码第三列；②尽管 Cys 和 Trp 物理性质不类似，但都由 Ser 衍生而来，它们的密码子共定位在 UGN 盒；③Sec 是 Ser 的另一个衍生物，由翻译前合成得到，它使用的 UGA 密码子，与 Cys 和 Trp 的密码子亦位于同一个 UGN 盒中（图 14.2）。这种系列关系非常明显，但这是统计学分析，还不能构成严格的证明。

对原理 3 的严格证明来自对原理 2 的证明。Gln 分配得到 CAA 和 CAG，这是因为携带这些密码子反密码子的 Glu-tRNA 通过翻译前合成转化成了 Gln-tRNA，而不是因为 Gln 和其密码子/反密码子间有任何立体化学相互作用、错误最小化或其他因素。同样，翻译前合成分配了 AAU 和 AAC 给 Asn，UGU 和 UGC 给 Cys，UGA 部分使用给 Sec。因此，对这三个氨基酸的密码子分配，生物合成是唯一的决定因素，因而证明了原理 3。

原理 4：由遗传密码编码的氨基酸是可突变的，这样早期密码扩充过程才可以加进 2 期氨基酸。

有些生命体内和细胞器中都广为人知地含有密码子分配的微小差异，说明了前 LUCA 生命在密码子分配方面的可突变性。可是，在全生命体系中 20 个标准氨基酸的编码和使用在过去 30 亿年内都从没有变异，令人怀疑它们到底能不能突变。既然 CET 提出，密码随着 2 期氨基酸的加入曾经大规模调整过，那么被编码氨基酸在进化过程中就必须能够自由突变。如果长期的实验尝试都不能使被编码的氨基酸突变，原理 4 就不正确了。因此，我们于 1983 年对枯草芽孢杆菌的遗传密码进行了突变的尝试。14.5 节的结果显示，4-氟-Trp 可以取代 Trp，甚至将 Trp 从标准氨基酸行列中除掉[59~61]，这证明了原理 4。

基于这 4 个原理的证实，CET 得到证明。

14.4 对全生命遗传密码的选择

在密码中有天文数字的可能排列，去编排不同氨基酸的密码子。虽然在很多核和细胞器系统，尤其是后生动物线粒体中，密码子的分配都有一些小变异的例子[62]，但是一般来说，全生命体系都基本上使用着同一套密码子，这意味着由 LUCA 阶段开始，生物界就再没有另类密码的存在了。密码子的排列可以产生约 2×10^{19}种不同的可能编码[28]。由于宇宙只有 15×10^{9} 年即 4.7×10^{17} s 的历史，要选出一个全生命遗传密码必须每秒钟通过自然选择淘汰掉 40 种使用另类编码

的生命，这本身是不可能的。也就意味着，编码限制与编码淘汰必须同时进行才能带来独一无二的全生命密码的出现。三种主要进化机制参与了这独一无二的密码的筛选。

（1）错误最小化可以指导遗传密码走向增强纠错功能的方向，该机制为全生命密码的产生贡献了 1×10^{-6} 的选择因子，即百万分之一的筛选[34]。

（2）立体化学相互作用并不依赖生命形式间的竞争去剔除编码。那些氨基酸-密码子/反密码子间缺乏相互作用的编码根本不能形成，该机制为全生命密码的产生贡献了 4×10^{-4} 的选择因子，即万分之四的筛选[43]。

（3）氨基酸生物合成的共进化，按照生物合成将氨基酸分组，大大限制了密码子分配的变异幅度，例如，一旦 UGN 盒被分配给 Ser，就只有由 Ser 生物合成衍生的氨基酸如 Cys、Trp 和 Sec 能够得到 UGN 盒中再分出来的密码子，这个机制为唯一编码的产生贡献了 10^{-11} 的选择因子，即一千亿分之一的筛选[28]。这个减少编码选择的模式和婚礼上将关系密切的客人安排到一起是类似的(专题 14.2)。

专题 14.2　婚宴上的座位

不同氨基酸密码子的分配也许能类比为婚宴上来宾的席位。有 20 个来宾，A、B……S、T，有 20 个座位。总共有 $p!\ (q!)^p$ 种不同的座位安排，p 是座位区域的数目，q 是每个区域的人数。在第一种座位安排下，将所有 20 个来宾都看做是单独的，用抽签来决定座位安排，因而 $p=1$，$q=20$。从而总共产生 2.4×10^{18} 种不同的座位安排。

若有 5 组比较亲密的来宾，每组 4 人——①A～D 为新娘的亲戚，②E～H 为新郎的亲戚，③I～L 为新娘的同事，④M～P 为新郎的同事，⑤Q～T 为邻居，那么第二种方案就是将座位分为 5 个区域，随机抽签将这些区域分配给 5 个组。每个区域内的 4 个座位随机分配给同组的 4 个人，这样，$p=5$，$q=4$，总共只产生 9.6×10^{8} 种座位安排。因而，由于第二种方案中的这种限制，不让不同组坐在同一区域内，就使得可能的座位安排减少为原先的 4×10^{-10} 倍。

同样，要将遗传密码中的 64 个密码子分配给 20 个氨基酸和终止信号，若没有任何分组，可能的分配数就约为 2×10^{19}。如果根据生物合成关系将 20 个氨基酸分组，可能的密码数就为 2×10^{8}，可选密码减少了 10^{11} 倍，极大有利于从所有可能的密码中选出唯一一种。

这三个机制一共能够提供 4×10^{-21} 的筛选，这足够从 2×10^{19} 种可能编码中选出一个独一的编码方式，也就足以解释全生命密码产生之谜。这三种机制的相对贡献为：

共进化∶错误最小化∶立体化学相互作用＝40 000 000∶400∶1

因此，共进化过程在生物界进化中能够成功筛选出一个独一性的全生命遗传密码起了绝对的主导作用。

14.5 遗传密码突变

很多氨基酸类似物都很容易进入原核和真核细胞的蛋白质，但它们都不能支持细胞的无限生长。为了检验 CET 中的原理 4，我们于 1983 年进行了一个对枯草芽孢杆菌 QB928（有 Trp 营养缺陷性，并且不能从 Trp 依赖中恢复）氨基酸编码的突变实验，目的是要将长期细胞生长中对 Trp 的利用替换为其类似物 4-氟-Trp（4FT）。经过两个突变阶段，QB928 产生了 LC33 细胞株（LC＝大菌落 4FT 细胞），它能够在 4FT 琼脂上形成长期生长的菌落（图 14.4）。尽管两个突变阶段可能含有多于两次突变，但要达到 4FT 长期使用所需要的突变次数仍然出乎意料的少。LC33 经历了另外两步突变后更产生了 HR15（HR＝高 4FT/Trp 生长比例），它在 4FT 中生长良好，在 Trp 中则不行，因此，在 LC33 中，Trp 被 4FT 取代了。而在 HR15 及其衍生细胞株中，如生长较快的 HR23 和依赖 Met 的 MR3，Trp 实际上是从氨基酸编码中被淘汰了，再也不能支持长期的细胞生长。这些依赖 4FT 的细胞株的生长现在都被变为类似物抑制剂的 Trp 抑制。但是，这些细胞亦可以通过回复突变重新获得利用 Trp 长期生长的能力（图 14.5）。LC33 可在 Trp 和 4FT 中生长，通过它的突变产生了 LC88 细胞株，它能在 Trp、4FT、5-氟-Trp（5FT）和 6-氟-Trp（6FT）（图 14.4）中生长，虽然 5FT 是细菌生长的一种强力抑制剂[59~61]。

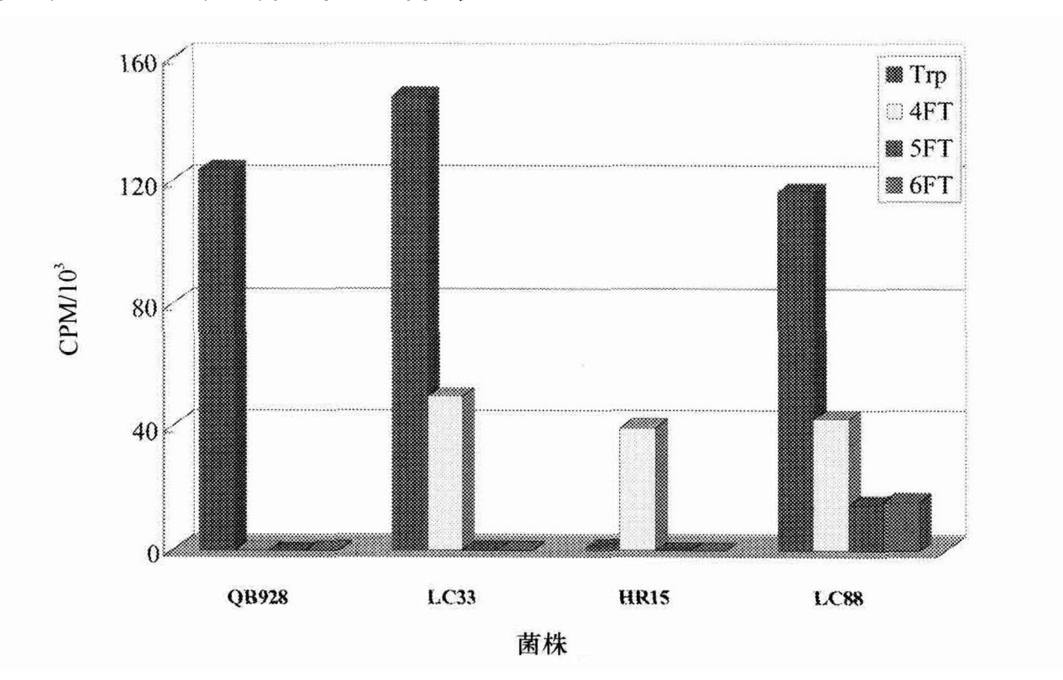


图 14.4 枯草芽孢杆菌菌株在 Trp 和 4-氟-Trp、5-氟-Trp 或 6-氟-Trp 上的生长情况，由［^{33}P］-磷酸标记，琼脂培养[61]（见彩插）。

在 HR15 中 Trp 不能再作为标准氨基酸的事实显示遗传密码进化是不限于加入新氨基酸的。有些氨基酸更可以被试用，发现不合适后被淘汰掉，这种淘汰

更换足以解释 α-氨基丁酸在编码中令人费解的缺席，该氨基酸很容易在大气中非生物合成出来。在表 9.1 的电火花合成中，生成的 α-氨基丁酸量比 Gly 和 Ala 少，但比 Glu 和 Asp 还多。它在现今遗传密码中的缺席可能是因为刚开始时为生命所采用，之后再被剔除。作为一个四碳的疏水氨基酸，它可能被 Val、Leu 和 Ile 这些具有更大烷烃链、疏水性更强的氨基酸所取代了。

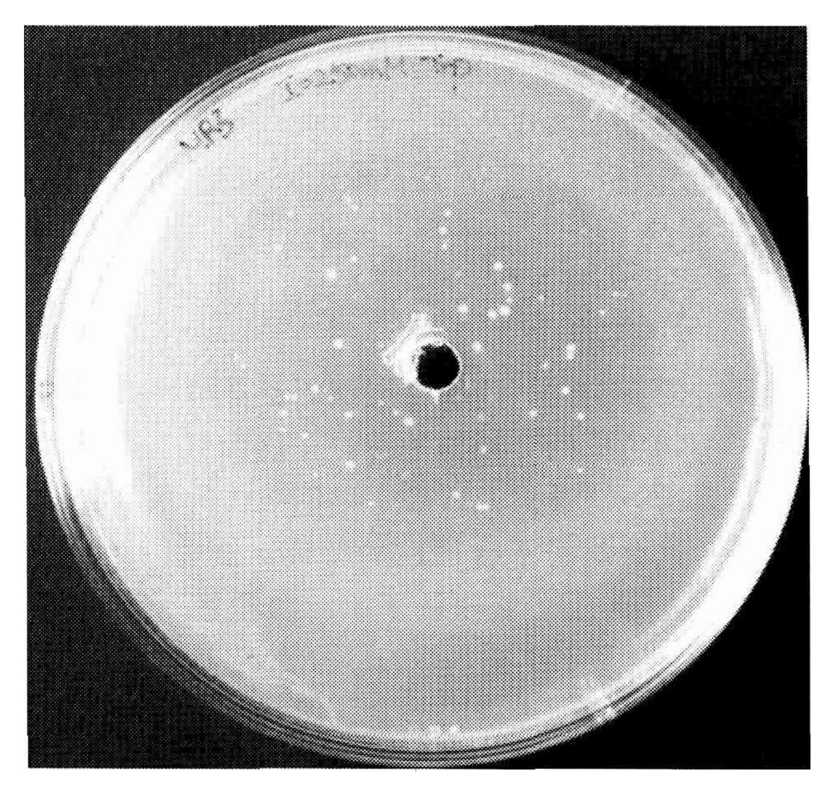


图 14.5　4FT 环境中，中央凹陷处的 Trp 抑制枯草芽孢杆菌 MR3 的生长。回复体(revertant)克隆重获在 Trp 中生长的能力，在干净的 Trp 抑制区域内可见。

用其氟代类似物将 Trp 从编码中替换或淘汰，需要的突变步骤相对很少，这说明在枯草芽孢杆菌蛋白质组的 12 625 个 Trp 残基中，只有少数 Trp 残基周围的蛋白质结构非常关键，Trp 若被 4FT、5FT 或 6FT 取代，立即带来功能缺陷和生长抑制。若这种结构上的限制通过突变放松了，那么 4FT、5FT 和 6FT 对 Trp 的取代就可以被接受，如在 LC88 里面。若结构限制更进一步发生突变，蛋白质不能利用 Trp 发挥功能，只可以利用 4FT 发挥，那么 Trp 就会失去支持细胞长期生长的能力而变成抑制类似物，如在 HR15 里面。

将 4FT 导入枯草芽孢杆菌氨基酸编码，是将遗传密码变异，从而改变标准氨基酸的第一个例子。它为共进化理论的原理 4 提供了决定性的支持，证实了即使从生物物种的最早分支开始，遗传密码被冻结，标准氨基酸被固定下来，事实上遗传密码仍然具有可变的本质。这一度冻结的密码的可变性的发现，引进了遗传密码变异的新纪元。变异途径有两条：让 4FT 在枯草芽孢杆菌中得到编码的

由上到下的基因组途径已被应用于大肠杆菌和大肠杆菌噬菌体[63~65]。而由下到上的途径中，可以使用功能隔离性 aaRS-tRNA 对，将一个非正常氨基酸加入到蛋白质序列的特定位置[66,67]。由于这种功能隔离性 aaRS-tRNA 只对相互起作用，而对细胞中其他 aaRS 和 tRNA 的作用很小[68]，它们的使用对细胞的生长影响不大。

由于氨基酸系统早已被确定，在所有生命体中都不变，所以具有新氨基酸编码的生命体被认为是新的生命形式[69]。新氨基酸的加入，使传统 20 个氨基酸编码的古老束缚被解开，增加了进化的自由度，为生物界写下了新的续集[70]。枯草芽孢杆菌菌株 HR15，用 4FT 而放弃 Trp 作为标准氨基酸，是合成生命的第一个例子。

14.6 结论

全生命遗传密码分两期构建。前期密码编码给环境中产生的 1 期氨基酸，后来经过扩充，加入了生物合成的 2 期氨基酸。在这个过程中，2 期氨基酸分配到原属 1 期氨基酸的密码子。这种扩充打破了早期前体的氨基酸密码子域，建立了一种结构奇特的编码，具有似乎不合理性的特征，例如，不同氨基酸得到 1～6 个密码子的不均等分配，Ser 的分裂密码子域，物理上不相似的 Cys 和 Trp 密码子位于同一个 UGN 盒，以及 Pyl 和 Sec 的部分密码子使用。虽然这些结构有些奇特，但扩充的主要目的，即促进编码的全套氨基酸侧链的化学多性能，是完全实现了。这些侧链具有多样性和平衡性，使得蛋白质催化能力强大，许多扩散控制酶表现突出，催化反应速率只受底物向酶活性位点的扩散速度限制，也使得多细胞生命得以发展，人类智慧、艺术和科学得以出现。在深度冻结了 30 亿年后，这个编码在 1983 年被实验改变，其内在的可变异性再一次使密码扩充成为事实，用人造遗传密码去产生人造生命形式。

在进化中，虽然基于生命体间竞争的自然选择最为重要，但有人提出，发展制约因素在进化中也起了重要作用，例如，霸王龙退化的前肢可能是其他方面的适应变化，如强大后肢的副产物[72,73]。独一全生命遗传密码的产生清晰体现了制约因素的作用：氨基酸生物合成关系对编码的制约，能从众多的可能密码中做出 $1/(2\times10^{19})$ 的选择，甚至比直接竞争性剔除使用另类密码的物种更为重要。在 *Chance and Necessity*（《偶然性与必然性》）一书中，Jacques Monod[74] 提出了生物学最重要的三个前沿挑战：生命起源问题、遗传密码起源之谜和中枢神经系统起源之谜。在这些方面，共进化理论的证明确认了前体氨基酸密码子域的再分裂，从而容纳生物合成的氨基酸，这是解开遗传密码结构起源之谜的关键。正如从一个国家的地图常常能看出它的历史一样，全生命遗传密码的结构正是遗传密码与氨基酸原始合成途径共同进化历史的永久碑铭。

参考文献

扩展阅读

1. Eigen M, Winkler-Oswatitsch R. Steps towards life. Oxford University Press, 1992.
2. Gesteland RF, Cech TR, Atkins JF eds. The RNA world 2nd ed. Cold Spring Harbor Laboratory Press, 1999.
3. Joyce GF. The antiquity of RNA-based evolution. Nature 2002; 418:214-221.
4. Lahav N. Biogenesis. Oxford University Press, 1999.
5. Wong JT. Coevolution theory of the genetic code at age thirty. BioEssays 2005; 27:416-425.

专题文献

6. Gamow G. Possible relation between deoxyribonucleic acid and protein structures. Nature 1954; 173:318.
7. Gilbert W. The RNA World. Nature 1886; 319:618.
8. Benner SA, Burgstaller P, Battersby TR et al. Did the RNA World exploit an expanded genetic alphabet? In: Gesteland RF, Cech TR, Atkins JF eds. The RNA World 2nd ed. Woodbury:Cold Spring Harbor Laboratory Press. 1999; 163-181.
9. Motorin Y, Grosjean H. Appendix 1. Chemical structures and classification of posttranscriptionally modified nucleosides in RNA. In: Grosjean H, Benne R, eds. Modification and Editing of RNA. ASM Press 1998; 543-549.
10. Auffinger P, Westhof E. Appendix 5: Location and distribution of modified nucleotides in tRNA. In: Grosjean H, Benne R, eds. Modification and Editing of RNA. ASM Press 1998; 569-571.
11. Mat WK, Xue H, Wong JT. The genomics of LUCA. Frontiers in Biosc 2008; 13:5605-5613.
12. Lane BG. Historical perspectives on RNA nucleotide modifications. In: Grosjean H, Benne R, eds. Modification and Editing of RNA. ASM Press 1998; 1-20.
13. Cech TR, Herschlag D. Group I ribozymes: substrate recognition, catalytic strategies and comparative mechanistic analysis. In: Eckstein F, Lilley CMJ, eds. Catalytic RNA. Heidelberg:Springer-Verlag, 1996: 1-17.
14. L'Huillier PJ. Efficacy of hammerhead ribozymes targeting a-lactalbumin transcripts: experiments in cells and transgenic mice. In: Eckstein F, Lilley CMJ, eds. Catalytic RNA. Heidelberg:Springer-Verlag ,1996: 284-300.
15. Wong JT, Xue H. Self-perfecting evolution of heteropolymer building blocks and sequences as the basis of life. In: Palyi G, Zucchi C, Caglioti L, eds. Fundamentals of Life. Paris:Elsevier, 2002: 473-494.
16. Wong JT. Origin of genetically encoded protein synthesis: a model based on selection for RNA peptidation. Orig Life Evol Biosph 1991; 21:165-176.
17. Zhang B, Cech TR. Peptide bond formation by in vitro selected ribozymes. Nature 1997; 390:96-100.
18. Illangasekare M, Sanchez G, Nickles T et al. Aminoacyl-RNA synthesis catalyzed by an RNA. Science 1995; 267:643–647.
19. Saito H, Kourouklis D, Suga H. An in vitro evolved precursor tRNA with aminoacylation activity. EMBO J 2001; 20:1797-1806.
20. Robertson MP, Knudsen SM, Ellington AD. In vitro selection of ribozymes dependent on peptides for activity. RNA 2004; 10:114-127.
21. Roth A, Breaker RR. An amino acid as a cofactor for a catalytic polynucleotide. Proc Natl Acad Sci USA 1998; 95:6027-6031.
22. Lautru S, Challis GL. Substrate recognition by nonribosomal peptide synthetase multi-enzymes. Microbiol 2004; 150:1629-1636.
23. Szathmary E. Coding coenzyme handles: a hypothesis for the origin of the genetic code. Proc Natl Acad Sci USA 1993; 90:9916-9920.
24. Schimmel P, Henderson B. Possible role of aminoacyl-RNA complexes in noncoded peptide synthesis and origin of coded synthesis. Proc Natl Acad Sci USA 1994; 91:11283-11286.
25. Schimmel P, Ribas de Pouplana L. Transfer RNA: from minihelix to genetic code. Cell 1995; 81:983–986.
26. Wong JT. A coevolution theory of the genetic code. Proc Natl Acad Sci USA 1975; 72:1909-1912.
27. Wong JT. Co-evolution of the genetic code and amino acid biosynthesis. Trends Biochem Sci 1981; 6:33–36.
28. Wong JT. The evolution of a universal genetic code. Proc Natl Acad Sci USA 1976; 73:2336-2340.
29. Li Y, Zhao Y, Hatfield S et al. Dipeptide seryl-histidine and related oligopeptides cleave DNA, protein and carboxyl ester. Bioorg Med Chem 2000; 8:2675-2680.
30. http://en.wikipedia.org/wiki/Posttranslational_modification.
31. Crick FHC. The origin of the genetic code. J Mol Biol 1968; 38:367-379.
32. Sonneborn TM. Degeneracy of the genetic code, extent, nature and genetic implications. In: Bryson V, Vogel HJ, eds. Evolving genes and proteins. New York: Academic Press 1965: 379-397.
33. Woese CR, Dugre DH, Dugre SA et al. On the fundamental nature and evolution of the genetic code. Cold Spring Harbour Symp Quant Biol 1966; 31:723-736.
34. Freeland SJ, Hurst LD. The genetic code is one in a million. J Mol Evol 1998; 47:238-248.
35. Freeland SJ, Wu T, Keulman N. The case for an error minimizing standard genetic code. Orig Life Evol Biosph 2003; 33:457-477.
36. Ardell DH, Sella G. No accident: genetic codes freeze in error-correcting patterns of the standard genetic code. Phil Trans R Soc London B 2002; 357:1625-1642.
37. Grantham R. Amino acid difference formula to help explain protein evolution. Science 1974; 185:862-864.
38. Wong JT. Role of minimization of chemical distances between amino acids in the evolution of the genetic code. Proc Natl Acad Sci USA 1980; 77:1083-1086.
39. Di Giulio M, Medugno M. The historical factor: the biosynthetic relationships between amino acids and their physicochemical properties in the origin of the genetic code. J Mol Evol 1998; 46:615-621.
40. Rodin S, Rodin A, Ohno S. The presence of codon-anticodon pairs in the acceptor stems of tRNAs. Proc Natl Acad Sci USA 1996; 93:4537-4542.
41. Lacey JC Jr, Wickramasinghe NSMD, Cook GW. Experimental studies on the origin of the genetic code and the process of protein synthesis: a review update. Orig Life Evol Biosp 1992; 22:243-275.
42. Jungck JR. The genetic code as a periodic table. J Mol Evol 1978; 11:211-224.
43. Knight R, Landweber LF, Yarus M. Tests of a stereochemical genetic code. In: Lapointe J, Brakier-Gingras L eds. Translation Mechanisms. Landes Bioscence 2003; 115-12.
44. Legiewicz M, Yarus M. A more complex isoleucine aptamer with a cognate triplet. J Biol Chem 2005; 280:19815-19822.
45. Hendry LB, Bransome ED Jr, Hulson MS et al. First approximation of a stereochemical rationale for the genetic code based on the topography and physicochemical properties of "cavities" constructed from models of DNA. Proc Natl Acad Sci USA 1981; 78:7440-7444.
46. Shimizu M. Molecular basis for the genetic code. J Mol Evol 1982; 18:297-303.
47. Seligman H, Amzallag GN. Chemical interactions between amino acid and RNA: multiplicity of the levels of specificity explains origin of the genetic code. Naturwissen 2002; 89:542-551.
48. Pelc SR. Correlation between coding triplets and amino acids. Nature 1965; 207:597-599.
49. Dillon LS. The origins of the genetic code. Bot Rev 1973; 39:301-345.
50. Wong JT, Bronskill PM. Inadequacy of prebiotic synthesis as origin of proteinous amino acids. J Mol Evol 1979; 13:115-125.
51. Wong JT. Question 6: Coevolution theory of the genetic code: a proven theory. Orig Life Evol Biosph 2007; 37:403-408.
52. Di Giulio M. The coevolution theory of the origin of the genetic code. Physics Life Rev 2004; 1:128-137.
53. Trifonov EN, Gabdank I, Barash D et al. Primordia vita. Deconvolution from modern sequences. Orig Life Evol Biosph 2006; 36:559-565.
54. Wong JT, Chen J, Mat WK et al. Polyphasic evidence delineating the root of life and roots of biological domains. Gene 2007; 403:39-52.
55. O'Donoghue P, Sethi A, Woese CR et al. The evolutionary history of Cys-tRNA^{Cys} formation. Proc Natl Acad Sci USA 2005; 102:19003-19008.
56. Sauerwald A, Zhu W, Major TA et al. RNA-dependent cysteine biosynthesis in archaea. Science 2005; 307:1969-1972.
57. Commans S, Bock A. Selenocysteine inserting tRNAs: an overview. FEMS Microb Rev 1999; 23:335-351.
58. Feng L, Sheppard K, Namgoong S et al. Aminoacyl-tRNA synthesis by pretranslational amino acid modification. RNA Biology 2004; 1:16-20.
59. Wong JT. Membership mutation of the genetic code: loss of fitness by tryptophan. Proc Natl Acad Sci USA 1983; 80:6303-6306.
60. Bronskill PM, Wong JT. Suppression of fluorescence of tryptophan residues in proteins by replacement with 4-fluorotryptophan. Biochem J 1988; 249:305-308.
61. Mat FWK, Xue H, Wong JT. Genetic encoding of 4-,5-,6-fluorotryptophans: Role of oligogenic barriers. Am Soc Microbiol 104th Meeting 2004; R-029.
62. Abascal F, Posada D, Knight RD et al. Parallel evolution of the genetic code in arthropod mitochondrial genomes. PLoS Biology 2006; 4:e127.
63. Bacher JM, Ellington AD. Selection and characterization of Escherichia coli variants capable of growth on otherwise toxic tryptophan analogues. J Bacteriol 2001; 183:5414-5425.

64. Bacher JM, Ellington AD. The directed evolution of organismic chemistry: unnatural amino acid incorporation. In: Lapointe J, Brakier-Gingras L eds, Translation Mechanisms. Landes Bioscences 2003; 80-94.
65. Bacher JM, Hughes RA, Wong JT et al. Evolving new genetic codes. Trends Ecol Evol 2004; 19:69-75.
66. Xie J, Schultz PG. Adding amino acids to the to the genetic repertoire. Curr Opin Chem Biol 2005; 9:5548-554.
67. Budisa N. Engineering the genetic code. Wiley-VCH. 2006.
68. Kwok Y, Wong JT. Evolutionary relationships between halobacterium cutirubrum and eukaryotes determined by use of aminoacyl-tRNA synthetases as phylogenetic probes. Can J Biochem 1980; 58:213-218.
69. Hesman T. Code breakers. Scientists are altering bacteria in a most fundamental way. Science News 2000; 157:360-362.
70. Cohen P. Life the sequel. New Scientist 2000; 167:32-36.
71. Wong JT. Kinetics of enzyme mechanisms. London:Academic Press, 1975: 200-201.
72. Gould SJ, Lewontin RC. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc Roy Soc London B 1979; 205:581-598.
73. Larson EJ. Evolution. New York:Modern Library, 2004: 280-281.
74. Monod J. Chance and Necessity. New York:Vintage Books, 1972: 138-148.

15 生命之根

王子晖

15.1 引言

所有活的生命体基本都利用全生命遗传密码，它们之间只有少量变异。自最后一个全生命共同祖先（last universal common ancester，LUCA）以来的所有生命种系都是这样。LUCA 是生命的根源。要理解 LUCA 产生的途径，以及由 LUCA 发展到现今生命的途径，关键是要知道 LUCA 的性质。

寻找生命之根的一个重要方法是根据蛋白质序列的旁系同源物（paralog），即由同一个基因复制产生的姐妹序列，但行使着不同的生化功能。这些姐妹序列起始时基本相同。随时间推移，它们越来越偏离，而偏离量是衡量它们之间相距多远的标准。以前利用旁系同源蛋白质蛋白树的交联对 LUCA 进行寻找，最初定位于小亚基核糖体 RNA（SSU rRNA）树上的细菌域（见 2.4 节）。但是，只有非常少数的蛋白质在三个生命领域都找得到旁系同源序列可供分析，而且那个时候人们也不知道，这种寻根方法充满了伪迹，如长枝吸引、同水平基因转移和变异饱和，它们极易被到仅仅基于少数物种去寻根的结果误导[5,6]。例如，早期对旁系同源物 ValRS-IleRS 树的分析[7]，它并不包含足够的物种来察觉伪迹的存在[8]，这对复杂的 IleRS 树尤其是个严重问题。因此，早期旁系同源物分析将 LUCA 定位细菌域这一点是不足取信的，也带来了悲观的观点，令人质疑生命之根是否真的能够找到[9,10]。

15.2 转移核酸

由于通过蛋白质旁系同源物对 LUCA 的寻找容易遇到伪迹的问题，分析核心便投向其他带有序列信息的生物多聚物。DNA 和核糖体 RNA 不做考虑，因为细胞中没有 DNA 或 rRNA 的旁系同源物，这样就只有 tRNA 了。tRNA 序列的一个限制因素是，它们通常都很短，只含约 75 个碱基，其中有些还是半保守的。因此，由任何一个特定 tRNA 得到的序列信息量都很小，但通过分析一个物种的整个 tRNA 组可以克服这种限制，因为一个单个活细胞基因组包含多于 30 个 tRNA 基因，这些基因总共包含了超过 2000 个碱基。

15.2.1　异受体间距

根据遗传密码的共进化理论（见 14.3 节），在密码发展期间，一些属于前体氨基酸的 tRNA 与其反密码子一起被转移给产物氨基酸。也就是说，原始 tRNA 及其转化后的拷贝之间的联系，会在同物种不同氨基酸受体 tRNA 间留下可探测到的序列相似性，这些异受体 tRNA，与携带同一氨基酸的同工 tRNA 不同。顺序比较揭示，许多异受体 tRNA 具有很高的序列同源性，尤其是在全生命 tRNA 系统发生树上的古菌域。来自嗜热菌敏捷气热古菌（*Aeropyrum pernix*，Ape——见图 15.1 中的物种缩写）的 $tRNA^{Phe}$-$tRNA^{Tyr}$对相互间仅在 4 个碱基的位置上有所不同。无论是在细菌的 Eco 或真核的 Ecu 中，同样一对 tRNA 间的区别都要更大一些（图 15.2）。这些发现说明，$tRNA^{Phe}$-$tRNA^{Tyr}$为旁系同源物，由一个原始基因复制而来。正如亲姐妹的基因比嫡系表亲更近，而嫡系表亲又比远房表亲更近等，这两个 tRNA 旁系同源物同样随着进化的过程渐行渐远，因此，比起 Ecu 和 Eco，Ape 离开生命之根的进化显然更少一些[11]。

细胞中，氨酰基-tRNA 合成酶（aaRS）需要精确识别其匹配的 tRNA，而仅把该 tRNA 与其氨基酸底物进行合成。如果把非匹配 tRNA 错认，一个错误的氨基酸就会进入蛋白质。这种识别过程依赖于 aaRS 对匹配 tRNA 上的特异核苷酸残基上的鉴别元件进行确认，而非匹配 tRNA 上是不具有这种鉴别元件的。aaRS-tRNA 复合体的晶体结构显示，aaRS 及其匹配 tRNA 在这些特别位置进行密切接触。要达到精确识别，通常需要多于三个残基作为鉴别元件。因此，$tRNA^{Phe}$和 $tRNA^{Tyr}$，自从它们分别接受 Phe 和 Tyr，成为旁系同源物开始，它们之间一开始就需要维持三个或更多碱基的差异来作为用以区分的鉴别元件。也就是说，Ape 的 $tRNA^{Phe}$和 $tRNA^{Tyr}$，仅有 4 个碱基不同，在过去约 36 亿年内它们的顺序明显地只经历了极少的变化（图 15.2）。在蛋白质中还没有发现这样极端保守的序列。

在脊椎动物的进化中，已知的最早脊椎动物是一种无颚鱼。现在它们中的许多都已经灭绝了，但有些如七鳃鳗和盲鳗仍存活的很好。从系统发生的观点看，七鳃鳗会比兔子更接近古老的脊椎动物，但这并不是说七鳃鳗是比兔子更古老的生命体，实际上它们都是现代动物，而它们的世系都能够追溯到起源于约 5 亿年前寒武纪末期的脊椎动物共同祖先。但是，由于七鳃鳗世系居留在其先祖栖息地的水域，它们的进化需求相对较小，解剖和生理方面的改变较小。相反，兔子来到了陆地，改为呼吸空气，并开始用 4 条腿奔跑。很自然地，兔子的外表和行为就不像早期的无颚动物了，因此，要了解最早脊椎动物的呼吸机制和大脑结构等，七鳃鳗是比兔子更理想的模型。同样，通过 tRNA 序列判断，Ape（太古生物，不指灵长类）世系无论怎样都比 Eco 和 Ecu 经历了更少的分子进化。因此，

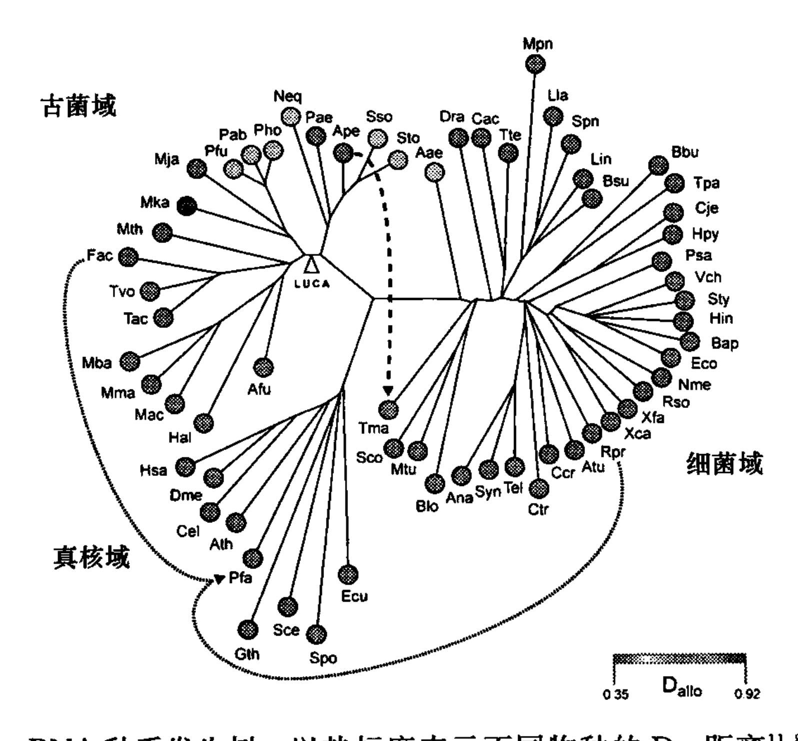


图 15.1 全生命 tRNA 种系发生树，以热标度表示不同物种的 D_{allo} 距离[11,8]。虚线显示接近 Tma 的 LBACA 由 Ape 的近亲形成；点虚线表示接近 Pfa 的 LECA 由与 Fac 接近的古菌宿主和类似 Rpr 的细菌通过胞内共生形成。物种名：泉古菌（ARCHAEA）。Ape 嗜热敏捷气热菌（*Aeropyrum pernix*）、Neq 骑行纳古菌（*Nanoarchaeum equitans*）、Pae 热棒菌（*Pyrobaculum aerophilum*）、Sso 硫叶菌（*Sulfolobus solfataricus*）、Sto（*Sulfolobus tokodaii*）；广古菌：Afu（*Archaeoglobus fulgidus*）、Fac 铁原体（*Ferroplasma acidarmanus*）、Hal 盐杆菌 NRC-1（*Halobacterium NRC-1*）、Mja 甲烷球菌（*Methanococcus jannaschii*）、Mka 甲烷火菌（*Methanopyrus kandleri*）、Mac（*Methanosarcina acetivorans*）、Mba 巴氏甲烷八叠球菌（*Methanosarcina barkeri*）、Mma 马氏甲烷八叠球菌（*Methanosarcina mazei*）、Mth（*Methanothermobacter thermautotrophicum*）、Neq 骑行纳古菌（*Nanoarchaeum equitans*）、Pab（*Pyrococcus abyssi*）、Pfu 激烈火球菌（*Pyrococcus furiosus*）、Pho（*Pyrococcus horikoshii*）、Tac 嗜酸热原体（*Thermoplasma acidophilum*）、Tvo（*Thermoplasma volcanium*）；细菌：Aae 超嗜热菌（*Aquifex aeolicus*）、Tma 海栖热袍菌（*Thermotoga maritima*）、Dra 耐辐射球菌（*Deinococcus radiodurans*）、Ctr 沙眼衣原体（*Chlamydia trachomatis*）、Bbu 博氏疏螺旋体（*Borrelia burgdorferi*）、Tpa 梅毒密螺旋体（*Treponema pallidum*）、Blo 长双歧杆菌（*Bifidobacterium longum*）、Mtu 结核分枝杆菌（*Mycobacterium tuberculosis*）、Sco 天蓝色链霉菌（*Streptomyces coelicolor*）、Ana 鱼腥藻（*Anabaena sp*）、Syn 胞藻 6803（*Synechocystis* 6803）、Tel（*Thermosynechococcus elongatus*）、Bsu 枯草芽孢杆菌（*Bacillus subtilis*）、Cac 丙酮丁醇梭菌（*Clostridium acetobutylicum*）、Lla 乳酸乳球菌（*Lactococcus lactis*）、Lin 无害李斯特菌（*Listeria innocua*）、Mpn 肺炎支原体（*Mycoplasma pneumoniae*）、Spn 肺炎链球菌（*Streptococcus pneumoniae*）、Tte 腾冲嗜热菌（*Thermoanaerobacter tengcongensis*）、Atu 根癌土壤杆菌（*Agrobacterium tumefaciens*）、Ccr 新月柄杆菌（*Caulobacter crescentus*）、Rpr 普氏立克次氏体（*Rickettsia prowazekii*）、Nme 脑膜炎双球菌（*Neisseria meningitidis*）、Rso 青枯病菌（*Ralstonia solanacearum*）、Bap（*Buchnera aphidicola*）、Eco 大肠杆菌（*Escherichia coli*）、Hin 流感嗜血杆菌（*Haemophilus influenzae*）、Psa 绿脓假单孢菌（*Pseudomonas aeruginosa*）、Sty 伤寒沙门氏菌（*Salmonella typhi*）、Vch 霍乱弧菌（*Vibrio cholerae*）、Xca 野油菜黄单胞菌（*Xanthomonas campestris*）、Xfa（*Xylella fastidios*）、Cje 空肠弯曲杆菌（*Campylobacter jejuni*）、Hpy 幽门螺杆菌（*Helicobacter pylori*）；真核：Ath 拟南芥（*Arabidopsis thaliana*）、Cel 秀丽隐杆线虫（*Caenorhabditis elegans*）、Dme 果蝇（*Drosophila melanogaster*）、Ecu 家兔脑胞内原虫（*Encephalitozoon cuniculi*）、Gth（*Guillardia theta*）、Hsa 智人（*Homo sapiens*）、Pfa 恶性疟原虫（*Plasmodium falciparum*）、Sce 酿酒酵母（*Saccharomyces cerevisiae*）、Spo 粟酒裂殖酵母（*Schizosaccharomyces pombe*）（见彩插）。



图 15.2　一些 tRNA-tRNA 对。其中不相同的碱基配对用阴影表示。

比起 Eco 和 Ecu，Ape 是更接近 LUCA 的模型。

任意一个游离细胞体的基因组都包含 20 类异受体 tRNA，而通过 190 对异受体每对间的平均基因距离，可以得到异受体间距（alloacceptor distance，D_{allo}），它的范围从最低的甲烷火菌 *Methanopyrus kandleri*（Mka）的 0.351，第二低的 Ape 的 0.402，到人类（Hsa）的 0.760 和最高的 Sce（即酵母）的 0.839。这个结果显示，Mka 的各种 tRNA 在序列空间上高度群集，相互间相似度很高。相反，Sce 的 tRNA 序列则分散得很远。这个结果可以用 tRNA 进化的集群发散（cluster dispersion）模型（图 15.3）来解释，tRNA 刚开始都聚集在图表中心的 *P* 处序列空间，又逐渐向外进化到未被占据的序列空间。在天空中，宇宙大爆炸从一个源点发散出星体，导致星体间的距离随时间不断增大。在由突变带来的 tRNA 顺序分散中，任意两个 tRNA 世系间的距离通常都随时间增大，如 tRNA A 和 tRNA B，但偶尔也会减小，如 tRNA C 和 tRNA D。对任何一个基因组，D_{allo}都能衡量基因组内各 tRNA 从 *P* 出发的平均进化发散距离：进化较慢的物种有较小的 D_{allo}，进化较快的则 D_{allo}较大。

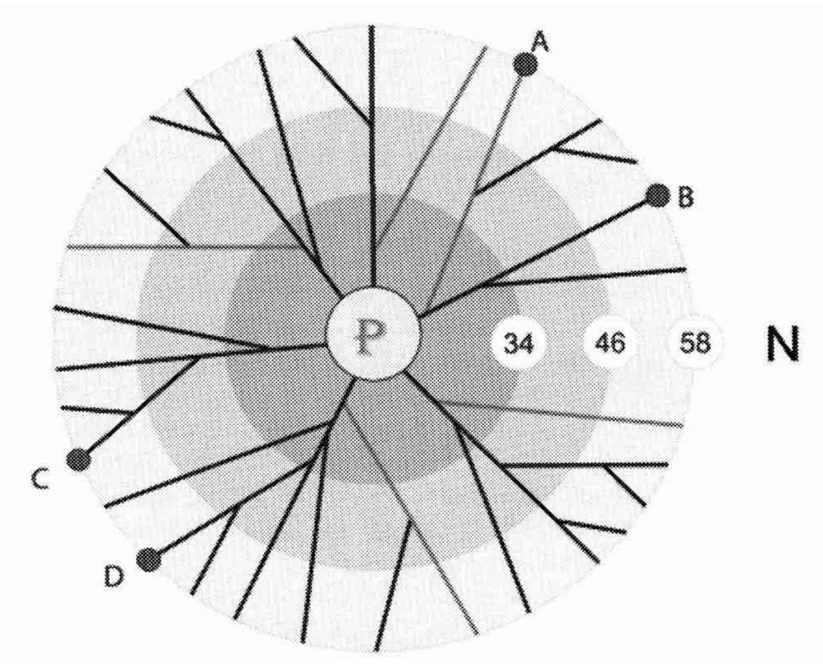


图 15.3　tRNA 进化的集群发散模型。分枝为基因复制，产生新的同工 tRNA（黑线）或异受体 tRNA（红线）。圆圈显示了 tRNA 基因组中不同阶段的 tRNA 基因数量（见彩插）。

基于这一点，LUCA 应该距离进化最少的基因组最近，这是通过它们的最小 D_{allo}距离确定的。图 15.1 表示了不同基因组在 tRNA 树上用热标度表达的 D_{allo}值。具有较低 D_{allo}的基因组并没有遍布整棵树，而是集中在较低的古菌分枝上，这说明 D_{allo}不是一个不稳定、无用的参数，而是具有很大用途的。由于广古菌 Mka 与泉古菌 Ape 代表了两个具有最低 D_{allo}的基因组，LUCA 应位于广古菌与泉古菌的联枝上[11]。而且，由于延伸 $tRNA^{Met}$和起始 $tRNA^{Met}$接受相同氨基酸，但功能不同，计算 D_{allo}时并未把它们看做异受体。当测量不同基因组间这两个 tRNA 的距离时，发现它们间最小的距离又属于 Mka[12]，因此，$tRNA^{Met}$的 D_{allo}距离和延伸-起始 $tRNA^{Met}$距离是将 LUCA 定位于最接近 Mka 处的第一和第二条证据（表 15.1）。第一条证据亦被约束性分析进一步证明，即古菌 tRNA 是病毒、真核生物和细菌 tRNA 的始祖[13]。

表 15.1　将 LUCA 定位于甲烷嗜热菌附近的一些证据[8]

序号	证据
1	tRNA 异受体间距
2	$tRNA^{Met}$起始因子-延长因子间距
3	反密码子利用
4	氨酰基-tRNA 合成酶间距
5	ValRS 的古细菌根
6	Mka 缺乏 GlnRS
7	Mka 缺乏 AsnRS
8	Mka 缺乏 CysRs
9	Mka 缺乏细胞色素
10	早期泉古菌-广古菌分离
11	Mka 处在古细菌域的较深分支
12	甲烷生成的原始性
13	厌氧生活方式的原始性
14	超嗜热性的原始性
15	适压性的原始性
16	嗜酸性的原始性
17	利用 CO_2 作电子受体
18	无机化能营养
19	热液口的优势
20	最简化的调节

全生命 tRNA 树确证了 SSU rRNA 树的三域结构（图 2.3），Woese 的三域推论是分子进化研究中卓越的成就[14,15]。但是 tRNA 树显示，在细菌域，革兰氏阳性菌相较于蓝藻处于更底层的分枝上，这与早期 SSU rRNA 树[16]相反，但与近代 SSU rRNA 树[17]和 SSU/LSU rRNA 树相同[18]。由于 SSU rRNA 没有旁系同源物，因而自身并不能为寻找生命之根提供帮助。早期将 SSU rRNA 树根定位于细菌域而不是古菌域[15]，只不过是用延伸蛋白因子及 ATP 酶旁系同源顺序测定，而各自只用了一个古菌物种，这对应付蛋白树的各种伪迹，是远远不足的。因此，LUCA 在 tRNA树（图 15.1）上古菌域内的定位，仅与这种用旁系同源蛋白质为 SSU rRNA 树测定根位的尝试结论不同，并不违背 SSU rRNA 树的三域结构。

早些时候，统计几何学对 tRNA 序列空间的分析表明，翻译过程起始于富含 GC 序列、长度小于 100 个核苷酸的 RNA 片段[19]。接近 LUCA 的 Mka，其 GC 含量达到 72.5%，这进一步证实了这个论断。图 15.4 显示，富含 GC 的共识 Mka tRNA 是先祖 tRNA 原型，即使大部分 Mka tRNA 及共识 Mka tRNA 在反密码子主干的 3′端不具有长的可变链，可变链在起源上可能是古老的，并可能在 LUCA tRNA 中存在[13]。在翻译出现之前，与 tRNA 原型相似的序列亦可能起了其他作用，如多肽化核酶的形成（见 14.2 节）、RNA 3′-可氨酰化结构[20]或作为线性 RNA 基因组 3′端复制起始的“基因组标记”[21]。早期 tRNA 基因的编码已在 13.5 节中讨论过。

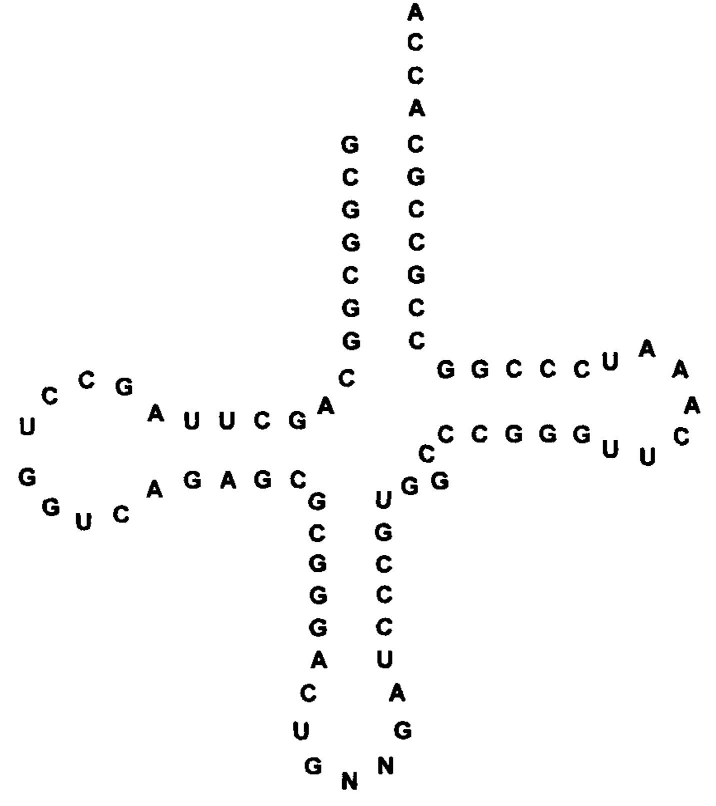


图 15.4　共识甲烷嗜热菌 tRNA 顺序。

15.2.2 反密码子的利用

基因组序列中找到的所有 tRNA 基因总体，显示了这个物种采用了哪一些反密码子，这种物种特异性的反密码子利用，为不同物种中如何实现遗传编码提供了大量有趣的信息。在遗传密码中，有 13 个标准 4 密码子盒，每个盒中的 4 个密码子要么分配给同一个氨基酸，如缬氨酸的 GUN 盒，要么平均分配给两个氨基酸，如天冬酰胺的 AAU-AAC 和赖氨酸的 AAA-AAG。在多数细菌和真核生物中，这 13 个盒是通过盒内 4 种反密码子的不同组合读取的（图 15.5）。比如，酵母（Sce）使用三个携带 3′G、U 或 C 的反密码子，来读取 UUN、CAN、AAN、GAN、AGN 和 GGN 盒；三个携带 3′A、U 或 C 的反密码子，来读取 GUN、UCN 和 ACN 盒；两个携带 3′G 或 U 的反密码子读取 CUN 盒；两个携带 3′A 或 U 的反密码子读取 CCN 和 GCN 盒；两个携带 3′A 或 C 的反密码子读取 CGN 盒。相反，没有古菌利用多于两种不同的反密码子组合。整个结果划出了明显区分，一方面是主要为反密码子多组合使用的细菌和真核生物，另一方面是主要为简单的、单组合使用的古菌[22]，如下所述。

		Mka	Tac*, Tma	Tvo	Fac	Sce	Bsu	Eco
UUU	*F*							
UUC	*F*	G	G	G	G	G	G	G
UUA	*L*	U	U	U	U	U	U	U
UUG	*L*		C	C	C	C	C	C
CUU	*L*				A			
CUC	*L*	G	G	G		G	G	G
CUA	*L*	U	U	U	U	U	U	U
CUG	*L*		C	C	C		C	C
GUU	*V*					A		
GUC	*V*	G	G	G	G		G	G
GUA	*V*	U	U	U	U	U	U	U
GUG	*V*		C	C	C	C		

		Mka	Tac*, Tma	Tvo	Fac	Sce	Bsu	Eco
UCU	*S*					A		
UCC	*S*	G	G	G	G		G	G
UCA	*S*	U	U	U	U	U	U	U
UCG	*S*		C	C	C	C		C
CCU	*P*					A		
CCC	*P*	G	G	G	G			G
CCA	*P*	U	U	U	U	U	U	U
CCG	*P*		C	C	C			C
ACU	*T*					A		
ACC	*T*	G	G	G	G		G	G
ACA	*T*	U	U	U	U	U	U	U
ACG	*T*		C	C	C	C		C
GCU	*A*					A		
GCC	*A*	G	G	G	G		G	G
GCA	*A*	U	U	U	U	U	U	U
GCG	*A*		C	C	C			

		Mka	Tac*, Tma	Tvo	Fac	Sce	Bsu	Eco
CAU	*H*							
CAC	*H*	G	G	G	G	G	G	G
CAA	*Q*	U	U	U	U	U	U	U
CAG	*Q*		C	C	C	C		C
AAU	*N*							
AAC	*N*	G	G	G	G	G	G	G
AAA	*K*	U	U	U	U	U	U	U
AAG	*K*		C	C	C	C		
GAU	*D*							
GAC	*D*	G	G	G	G	G	G	G
GAA	*E*	U	U	U	U	U	U	U
GAU	*E*		C	C	C	C		

		Mka	Tac*, Tma	Tvo	Fac	Sce	Bsu	Eco
CGU	*R*					A	A	A
CGC	*R*	G	G	G	G			
CGA	*R*	U	U	U	U			
CGG	*R*		C	C	C	C	C	C
AGU	*S*							
AGC	*S*	G	G	G	G	G	G	G
AGA	*R*	U	U	U	U	U	U	U
AGG	*R*		C	C	C	C	C	C
GGU	*G*							
GGC	*G*	G	G	G	G	G	G	G
GGA	*G*	U	U	U	U	U	U	U
GGG	*G*		C		C	C		C

图 15.5 反密码子利用[22]（以下系列古菌共用 Tac 模式：Ape、Pae、Sso、Sto、Afu、Hal、Mac、Mba、Mma、Pab、Pfu 和 Pho）（见彩插）。

细菌，34 个非寄生物种：两种利用 5 个或更多组合，7 种利用 4 个组合，19 种利用 3 个组合，5 种利用 2 个组合，只有 Tma 利用单一的 GNN＋UNN＋CNN 组合。

真核生物，7 个非寄生物种：1 种利用 5 个组合，1 种利用 4 个组合，3 种利用 3 个组合，2 种利用 2 个组合。

古菌，18 个非寄生物种：13 种都利用单一的 GNN＋UNN＋CNN 三反密码子

组合。Mka 则更为简单，利用单一的 GNN＋UNN 两种反密码子组合，这与其接近 LUCA 是吻合的。Mja、Mth 和 Tvo 同时利用这两类组合。Fac 在其标准盒的 12/13都仅仅利用 GNN＋UNN＋CNN，而在亮氨酸盒中利用 ANN＋UNN＋CNN。

进化通常由简单向复杂进行，因此比起细菌和真核生物，较简单的古菌反密码子使用方式表明古菌更为原始（表 15.1，第 3 行）。另外，有假设基于核糖体形态提议，应当将“Eocyte”域从古菌域分割出来[23]。但是，所有 4 种泉古菌和多数广古菌都利用单一 GNN＋UNN＋CNN 组合，有力地证明了古菌域的划一性。

虽然古菌是与真核生物及细菌种系独立发生的，迄今为止，只有极少数古菌特征被发现是其他两个域的生物所不具备的，例如，古菌利用醚-脂，而细菌和真核生物利用酯-脂，因此，对反密码子的简单或复杂使用是古菌与细菌-真核生物间稀有的基本分子生物特征的区别。该区别的进化意义是什么呢？Knud Nierhaus 提出了一个可能的答案（个人交流），这可能与核糖体延长因子 LepA 相关，这个因子是最为高度保守的蛋白质之一，存在于所有细菌和几乎所有真核生物中，但在古菌中没有。它能够令翻译过程中进行回复易位，防止核糖体停止运作，并增强核糖体对离子浓度改变的耐受性[24]。因此，细菌有了 Lep A，对内环境的改变就可能比古菌更能适应，而具有更好适应多种生态环境的能力。为了应对内环境的改变，密码子-反密码子的碱基配对强度也可能需要适当调整，于是有了细菌和真核生物的反密码子多组合使用。出于这种可能，由古菌域中形成细菌域，可能是有如 Lep A 这样的适应性进化的结果，使细菌远比古菌更易进入包括人体在内的新生态环境。这有助于解释细菌为什么在生态环境中广泛分布，而古菌相对受限制于极端环境，也有助于解释为什么人类有许多细菌引起的感染，而极少或没有由古菌引起的感染。

15.3 蛋白质

通过化石，古生物学研究揭示了寒武纪以来地球上生命体的恢弘全景。在更早的 6 亿年前，前寒武纪时期，艾迪卡拉动物群出现在澳大利亚的阿德莱德市附近，包含如水母和分节蠕虫的无壳标本。更为古老的是，能够用显微镜在岩石中发现微生物化石的独特均一性印迹，并通过蛋白质和核酸序列追溯分子进化过程。蛋白质如细胞色素 c、组蛋白和血红素为生物进化提供了无法估量的信息。但是，要追溯 30 亿年以前的 LUCA，蛋白质可能进化得太远，并有同水平基因转移（HGT）等干扰带来的问题。例如，在细菌域的 RNA 聚合酶 β 和 β' 亚基树上，产液菌位于远离分枝根部的位置，但若从树上除去两种类菌质体，则产液菌就会接近根部[25]。同样，多种蛋白质的结合树显示，最早的细菌群应是螺旋体或嗜热菌，这只取决于是否从树上除去某些物种[26]。

要从蛋白质序列中提取系统发生信息，有两种方法可以减少伪迹。通过成对比较可以估计出旁系同源蛋白质间的序列差距，而并不需要树的构建。另外，树的构建若能够囊括大量物种，也会帮助伪迹的暴露，这将在以下部分说明。

15.3.1 合成酶距离

BLASTP 算法可以用来估计出蛋白质之间的遗传距离[27]。将它应用于 tRNA 树上任意基因组的 20 种氨基酸 tRNA 合成酶（aaRS）的 190 对，得到 190 个比特（bit）值。若发现树上任意一个基因组内，两个 aaRS 间 bit 值大于 60，那么就可以拟定这两个 aaRS 可能是通过基因复制形成的旁系同源物。以这个标准，190 对 aaRS 中有 10 对可能是旁系同源物[12]。

在 tRNA 树上的基因组中，Mka 中的 ValRS-IleRS 这一对具有最高 bit 值，达到 506.5（表 15.2），毫无疑问，是旁系同源物。另外 9 对在同一个基因组中的 bit 值也都有大于 60 的例子，从 ThrRS-GlyRS 的 66.2 一路往上。最高的 bit 值并不总是在同一个基因组中出现，即使 10 对中有 5 对的最高分值都出现在 Mka 中。将任一基因组的这 10 个 bit 值求平均值，就得到了它的 Q_{ARS}商，可以用来衡量其 aaRS 旁系同源物在如今的相似度。图 15.6 用热标度显示了 tRNA 树上不同基因组的 Q_{ARS}。三个最高值分别为 Mka、Mth 和 Mja 的 138.5、119.3 和 115.2，远远超过了 Bsu 的 88.2、Eco 的 60.4 和人类（Hsa）的 40.9，显示出物种间存在明显的梯度。在 aaRS 基因型上，由于 Mka 具有最高的 Q_{ARS}分数，它应为图 15.6 树上进化最为缓慢的物种（表 15.1，第 4 行），正如图 15.1 所示的 tRNA 基因型树中，它具有最低的 D_{allo}值，从而也能显示其进化最为缓慢。但是，有些其他物种在这两棵树上表现了显著区别。Mth 和 Mja 的 tRNA 基因型进化得相当快，但 aaRS 基因型进化缓慢。然而，泉古菌的 tRNA 进化缓慢，但 aaRS 进化较快。

表 15.2 氨酰 tRNA 合成酶的可能旁系同源对[12]

aaRS 对	最大比特值	最匹配	次匹配	第三匹配
ValRS-IleRS	506.5	Mka	Mth	Mja
ValRS-LeuRS	232.3	Tma	Hal	Mka
LeuRs-IleRS	202.6	Pab	Pfu	Pho
MetRS-LeuRS	94.4	Tte	Cac	Bsu
ThrRS-ProRs	91.7	Mka	Pab	Sco
MetRS-ValRS	82.4	Cac	Lla	Tte
IleRS-MetRS	75.1	Mka	Lla	Aae
TrpRS-TyrRS	75.1	Mka	Sso	Sto
SerRS-ProRS	67.4	Lla	Spn	Tte
ThrRS-GlyRS	66.2	Mka	Pho	Pfu
平均（Q_{ARS}）		Mka 138.5	Mth 119.3	Mja 115.2

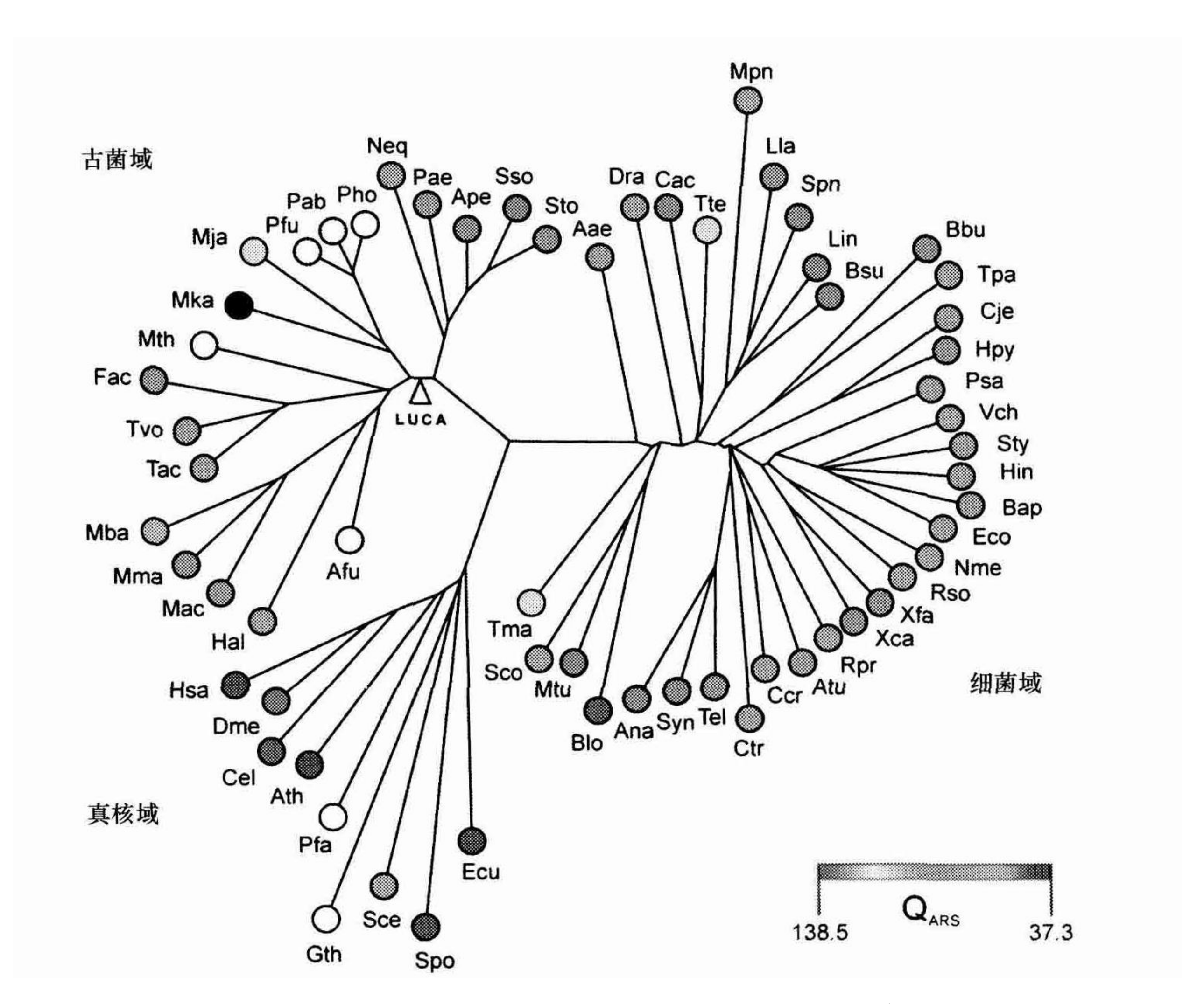


图 15.6　全生命 tRNA 树，以热标度显示了不同物种的 Q_{ARS}分数[12]（见彩插）。

15.3.2　ValRS 树的古菌鼻祖

由于在所有 aaRS 旁系同源物中，ValRS-IleRS 这一对具有最高的比特值，因此，若不存在过度的混乱现象的话，这两个蛋白质序列应当适合旁系同源物分析。为了检测混乱现象，采用了图 15.7tRNA 树上所有基因组中的 ValRS 和 IleRS序列来建立旁系同源树。古菌、细菌和真核生物的 IleRS 序列并不清晰分组，使得树的中间集群包含了三个域的序列，因此无法利用 ValRS 序列作为外群（outgroup）来寻找 IleRS 树的根。ValRS 序列比较有秩序，可以分得开不同领域各自的集群，即使 Blo 和 Rpr 的细菌序列跑到古菌集群中了。因此，利用 IleRS 序列作为外群来寻找 ValRS 树的根相对有效。红线上显示，IleRS 外群加入到 ValRS 古菌-细菌-真核三方连接处（J 点）的古菌那一边，因此将 ValRS 树的根定位于古菌域内（表 15.1，第 5 行）[8]。

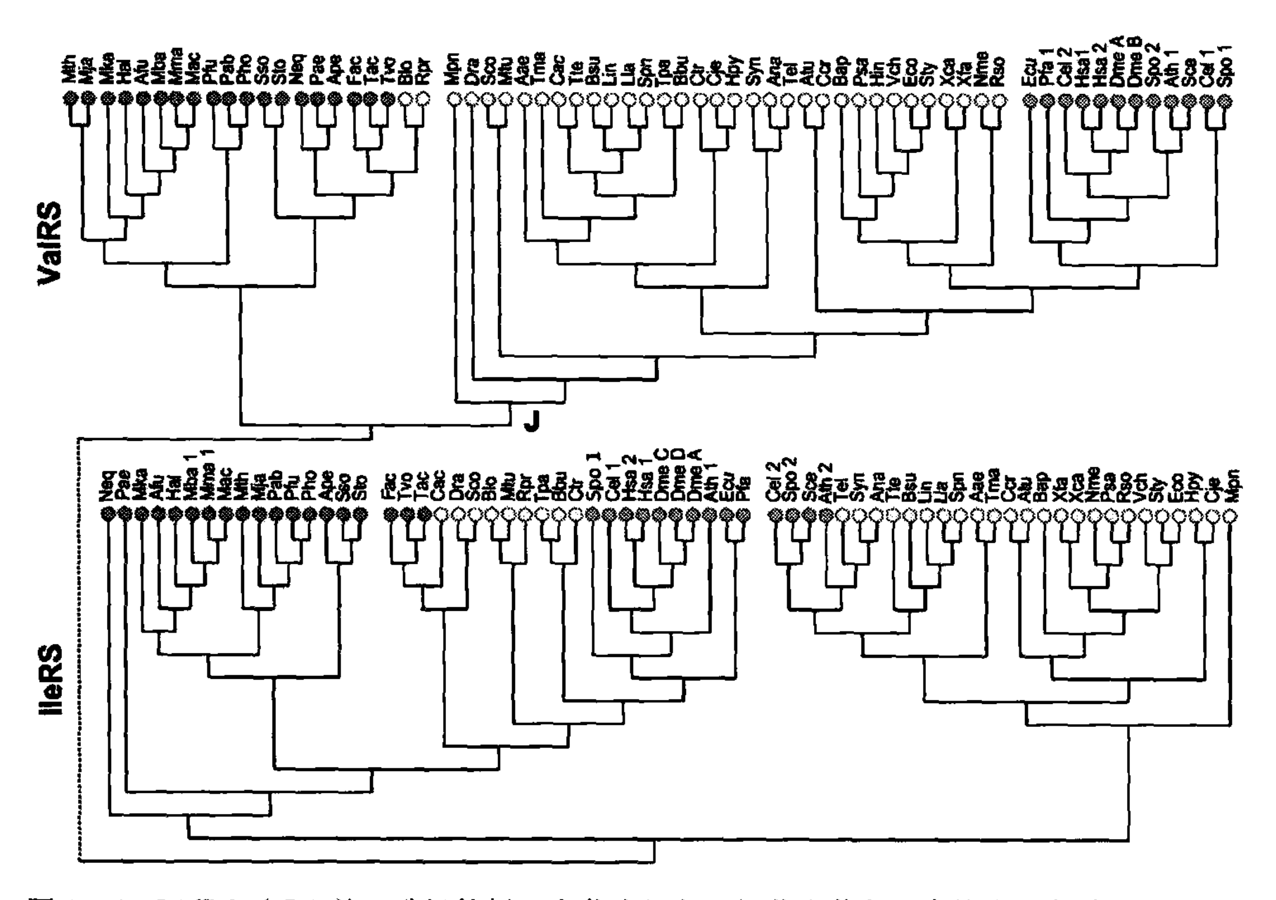


图 15.7 ValRS-IleRS 的一致极简树（古菌为红色，细菌为黄色，真核为蓝色）[8]（见彩插）。

15.3.3 欠缺的基因

在现今的生命体系中，一些物种中的 2 期氨基酸 Gln、Asn 和 Cys 分别通过 GlnRS、AsnRS 和 CysRS 进入蛋白质，但在另一些物种中则分别利用 Glu-tRNA、Asn-tRNA 和 *O*-磷酸-Ser-tRNA 通过翻译前（pretran）合成进入。在遗传密码进化中，pretran 合成的利用早于 aaRS 的利用（见 14.3 节）。因此，不用 GlnRS、AsnRS 或 CysRS，而用 pretran 合成来将 Gln、Asn 和 Cys 组装到蛋白质上是更原始的操作方式，因而没有进化离开 LUCA 多远的物种很可能会欠缺这三种 aaRS。这三种 aaRS 都不存在于 Mka 基因组中，更进一步证明了 Mka 与 LUCA 的近似（表 15.1，第 6～8 行）。

细胞色素为含有原血红素的蛋白质，广泛参与线粒体、叶绿体、硫酸盐还原生物，甚至属厌氧产甲烷菌的甲烷八叠球菌（*Methanosarcina*）中的电子转移，因此，细胞色素基因在 Mka、Mth、Mja、Pfu、Pab 和 Pho 基因组中的缺失是令人惊异的，这些物种在 tRNA 树上聚在一起。由于 Mka、Pfu、Pab 和 Pho 都具有很低的 D_{allo} 距离，而 Mka、Mth 和 Mja 具有最高的 Q_{ARS} 商，这些欠缺细胞色素的物种被确认为分子进化中异常保守的远古六菌（ancient six）。因为 Mka、Mth 和 Mja 利用 H_2 和 CO_2 合成甲烷，而 Pfu、Pab 和 Pho 则代谢产生 H_2 和

CO_2，所以它们的细胞色素缺失显然并不起因于共同代谢形式，而是表明它们与缺乏细胞色素的 LUCA 关系很近，由此，Mka 的细胞色素缺失写下了表 15.1 的第 9 行[8]。

15.3.4 原始蛋白

由 32 种蛋白质建立的复合系统发生树，表现了蛋白质进化的时间表，指出广古菌-泉古菌的分离是最古老的生物界事件之一，Mka 是最早的古菌分枝，而生物甲烷合成出现于 3.8～4.1 Gya 之前[28]（表 15.1，第 10～12 行）。

产氧光合作用带来的大气氧气积累增强了很多生命体中蛋白质组的适应调整。在游离的活生命体中，Tma、Mka 和 Mth 的后氧蛋白质组适应度最低，也就是说，Mka 是极为保守的生命体，保留了 LUCA 的厌氧特性[29]，而重建的原始蛋白的氨基酸组成也说明了这一点（表 15.1，第 13 行）[30]。原始蛋白重建及氨基酸密码结构的结合，亦表明 LUCA 是一种超嗜热菌，一种生活在海洋深处的适压菌，也是适酸菌，与今天的 Mka 非常类似[31～33]。古菌和细菌亦各自表现出超嗜热和原始性的关联[34]。同时，通过克隆和表达得到的古老延长因子具热稳定性，这说明很久以前的原始生命体生活在高温当中[35,36]。这些发现都揭示了 Mka 和 LUCA 的共性，给出表 15.1 中的第 14～16 行。

15.4 代谢作用

生命伊始，大量的电子受体未必是必需的，因为生命可以在一个小而适宜的孵育器中发展，但是后来它们应该是必需的，否则 LUCA 世系难以大量繁殖，这令 LUCA 足以击败所有竞争者，成为未来所有生命体的共同祖先。早期地球上大量的电子受体为 CO_2 及较少的硫酸盐[37,38]。因此，类似 Mka 这样的甲烷菌利用 CO_2 作为电子受体进行电子转移，就很有条件成为 LUCA（表 15.1，第 17 行）。像甲烷菌一样，利用 CO_2 作为电子受体，进行化能无机营养生命，更可以摆脱对有机能量底物的依赖，而有利于大规模繁殖（表 15.1，第 18 行）。Mka 生活的热液口是 LUCA 生长的适宜环境，这是因为热液口的产甲烷作用使得化能无机自养成为可能，除去大量 CO_2 作为电子受体，还增加了大量可作为电子供体的 H_2。同样，热液口的高温会迫使生命发展稳定的生化能力，如完美的遗传密码和氨基酸字母表及 DNA 信息机制（表 15.1，第 19 行）。Mka 最低限度的代谢调节机制[39]与其原始性也是协调一致的（表 15.1，第 20 行）。

15.5 细菌与真核生物的始祖

15.5.1 全细菌共同祖先

由于 LUCA 是一种古菌，所以它同样也是最后一个全古菌共同祖先

(LACA)，产生古菌领域。但是，最后一个全细菌共同祖先（Last Bacterial Common Ancestor，LBACA）及最后一个全真核共同祖先（Last Eukaryotic Common Ancestors，LECA）仍需要确定。在这个问题上 tRNA 序列同样是很有用的。因为所有的细菌都是通过 LBACA 从古菌演化而来，所以与 LBACA 相近的细菌，其 tRNA 很可能就比其他细菌的 tRNA 保留了更多与古菌的相似性。为了研究这种可能，我们将图 15.1 中所有古菌 tRNA 与所有细菌 tRNA 进行了单独比较。结果显示，在多于 1.2 M 个古菌-细菌 tRNA 对中，有 6 个高似（hi-sim）对具有高的序列相似性，其遗传距离均小于 0.15（表 15.3）。这 6 个高似对包括细菌域中 Tma 的三种、Aae 的两种，其中距离最小的两对均来自 Tma。在古菌域方面，6 个高似对中 Ape 占了三个，其中包括两个距离最小的。也就是说，LBACA 在细菌域距 Tma 最近，距 Aae 次之，而它的古菌祖先则与泉古菌 Ape 最近（图 15.1）。这同样符合 tRNA 树的拓扑结构，它显示在细菌域中 Aae 和 Tma 是两个最底层的分枝，而细菌域是由 LUCA 的泉古菌一侧发展起来的，并不是广古菌一侧。况且，Tma 利用单一的 GNN＋UNN＋CNN 反密码子去读取其全部标准密码子盒，这和 Ape 及绝大部分古菌（图 15.5）是一致的，有力地支持了它与 LBACA 的近似。相反，Aae 有 8 个盒利用 GNN＋UNN＋CNN、4 个盒利用 GNN＋UNN、1 个盒利用 ANN＋CNN，其三种反密码子组合方式属典型的细菌，与古菌特征完全不同。

表 15.3　高似和超高似跨领域 tRNA 对[8]

古细胞-细菌对	间距
Ape（Ile）cau-Tma（Ile）cau	0.104
Ape（Ile）gau-Tma（Ile）cau	0.104
Mja（Asn）guu-Mtu（Lys）uuu	0.120
Ape（Thr）cgu-Aae（Thr）cgu	0.122
Afu（Ser）gga-Tma（Ser）gga	0.139
Sso（Thr）cgu-Aae（Thr）cgu	0.139
真核生物-古细菌对	**间距**
Pfa（Phe）gaa-Fac（Phe）gaa	0.088
Pfa（Phe）gaa-Tac（Phe）gaa	0.104
Pfa（Phe）gaa-Tvo（Phe）gaa	0.104
Hsa（Phe）gaa-Fac（Phe）gaa	0.137
Pfa（Leu）uaa-Sto（Leu）caa	0.139

值得一提的是，Tma 是厌氧的，而 Aae 是好氧的。因此有必要检测与 Ape

相关的厌氧及微量需氧物种，以及与 Tma 相关的微量需氧和好氧物种，看看它们是否比 Ape 和 Tma 更接近古菌-细菌中间物种。

15.5.2　全真核共同祖先

将所有真核 tRNA 与所有古菌和细菌 tRNA 单独比较，结果显示，在多于 0.98 M 的真核-细菌对及 0.58 M 的真核-古菌 tRNA 对中，有高似对，还有遗传距离小于 0.10 的超高似对。检测到的 8 个超高似真核-细菌对都包括植物 Ath 作为真核方，蓝藻细菌 Tel、Ana 和 Syn 作为细菌方[8]。因为蓝藻细菌与叶绿体的前体关系密切，Ath 和蓝藻细菌间的高相似度很可能是因为来源于叶绿体基因组的蓝藻 tRNA 基因迁移到了植物细胞的核基因组中。

检测到的一个超高似及 4 个高似真核-古菌对，包含 Phe 和 Leu 的 tRNA，在真核域主要是疟原虫（Pfa），在古菌域主要是热原体属的 Fac（2x）、Tac（1x）和 Tvo（1x）（表 15.3）。唯一的超高似对为 Pfa（Phe）gaa-Fac（Phe）gaa，接受 Phe 并携有 GAA 反密码子（图 15.8）。这显示 LECA 可能起源于一种与铁原体（ferroplasma，Fac）相关的古菌宿主，与 Rpr 及可能还有其他细菌通过胞内共生作用形成。Fac 生活在富含金属硫化物的环境中[40,41]（图 15.1）。这种说法有以下证据支持。

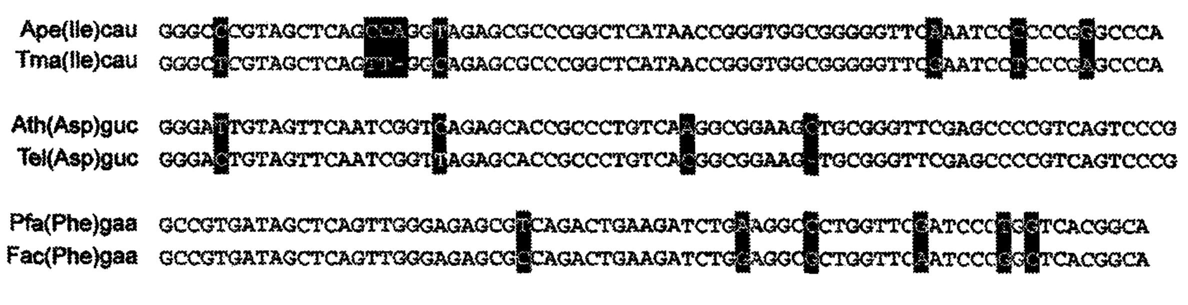


图 15.8　一些高似和超高似跨区域 tRNA 对。其中不同的碱基配对用阴影标出[8]。

（1）真核生物中 ANN＋UNN＋CNN 的反密码子组合最为普遍，它们利用这组合去读取 67％的氨基酸家族密码盒。相反，除了古菌 Fac（图 15.5）及细菌 Lla[22] 中的 CUN 盒，任何古菌或细菌都不使用这个组合，也就是说，该组合的使用是通过 Fac 相关寄主的胞内共生作用引入真核生物的。

（2）热原体缺乏细胞壁，这使得它适于作为寄主吞入细菌进行胞内共生[42]，而铁原体也具有这种良好特性。

（3）热原体的生化特性使它适于在富含硫化物的环境中进行胞内共生，从而作为真核生物的古菌前体[43]，这对铁原体同样也适用。

（4）超级树的种系发生信号剥离分析方法，亦揭示了真核-热原体属的关系[44]。

(5) 热原体为嗜热菌，其最佳生长温度为 55～59℃，而铁原体为嗜温菌，其最佳生长温度为 37℃[40,41]。因为没有嗜热的真核生物，而真核生物非常适宜于在 37℃生长，并且许多世代之后人体的温度也是 37℃，这支持嗜温铁原体的胞内共生，而不是嗜热热原体的胞内共生。

有很多证据表明，一种与立克次氏体（RPr）类似的 α-蛋白细菌通过胞内共生形成了真核生物线粒体的前体。Fac 成为真核生物祖先的胞内共生现象与立克次氏体成为线粒体的共生现象可能是一回事，但是也不排除是分开的，参与共生的细菌也可能相同或不同[45~47]。

真核-古菌的 Phe 受体（gaa）和 Leu 受体（uaa、caa）tRNA 具有高度的相似性，因此，通过原始真核生物及铁原体相关物种追踪这些 tRNA 序列，将能够促进对这些生命体系统发生过程的理解，并进一步确定它们与 LECA 的关系。

15.6 LUCA 基因组

15.6.1 非最小基因组

远古六菌——Mka、Mth、Mja、Pfu、Pab 和 Pho 的基因组缺乏细胞色素基因，显示出 LUCA 也同样缺乏这些基因。同样地，在所有这 6 个古菌基因组中都存在的基因也可以为 LUCA 基因组定义。但是，Mka-Mth-Mja 和 Pab-Pfu-Pho 这两组具有不同的能量代谢模式，因此，这两组共有的基因中就不会包含任意一方特有的能量代谢基因。若没有任何能量代谢基因，LUCA 就不可能存活，因此 LUCA 必须具备这方面的基因。由于 LUCA 最类似于超嗜热的甲烷菌，所以 Mka、Mth 和 Mja 共有的产甲烷基因可以代表 LUCA 的可能能量代谢基因（附录 15.1）[48]。

人们将蛋白质家族进行了直系同源聚类（COG）[49]。远古六菌共有的 COG 加上 Mka、Mth 和 Mja 共有的产甲烷 COG 使 LUCA COG 总数达到 561 个。这 561 个 COG 给出了 LUCA 蛋白质组的偏低估计。比如，Mka 及最像 LUCA 的物种包含 *SelD*（Mk1369）和 *SelA*（Mk0620）基因，用来为 Cys-tRNA 编码的翻译前合成，但远古六菌中的一些成员已经放弃了这个途径。因此这两个 COG 就不包含在这 561 个 COG 中。同时，在 tRNA 树上，LUCA 定位于广古菌和泉古菌之间。由于远古六菌都是原始的广古菌，所以它们共有的 COG 可能包含一些原来 LUCA 没有的基因，是后来才加入最早的广古菌后代中的。考虑到这种可能，附录 15.1 也显示了较小的一套 424 个 COG，不单是远古六菌所共有，同时也是加入 Ape 和 Pae 这两种具有最低 D_{allo}距离的泉古菌的远古八菌所共有。这 424 个 COG 定义了最小的 LUCA 蛋白质组。将其与 Mka 基因组里，39 个结构 RNA 基因合并产生了具有 463 个基因的最小 LUCA 基因组，这远远小于包含

1731个基因的Mka基因组，因此，要么在LUCA和Mka间基因组的大小曾经历了极大发展，要么这463个基因实际上低估了LUCA基因组，也可能两者皆是。有一种根据祖先状态的估计，认为LUCA基因组会大于1000[50]，这与463个基因的最小估计大致上是符合的。

LUCA到底是一种原始的前基因组合（progenote）还是具有成熟基因组的生命体，这是进化中悬而未决的关键问题之一[51]。模型Ⅰ～Ⅵ给出了能满足生命需要的最小蛋白质组的大小，包含150～340个蛋白质编码基因。

Ⅰ.“柔弱”生命中的基因（见2.7节）：150个基因

Ⅱ. Carsonella ruddii的最小基因组[52]：182个基因

Ⅲ.假设的最小细胞中的基因[53]：200个基因

Ⅳ.由生殖道支原体和流感嗜血杆菌基因组相比较推出的最小蛋白质组[54]：256个基因

Ⅴ.通过删除确定的枯草芽孢杆菌的必需基因[55]：271个基因

Ⅵ.细菌的核心蛋白质基因[56]：340个基因

LUCA具有424个COG的最小蛋白质组远远超过了Ⅰ～Ⅵ给出的数目，因此在LUCA基因组中一定还有一些“非最小”基因存在。例如，用于RNA加工和修饰的A组COG，用于染色质结构和动力的B组，用于细胞周期控制、细胞分裂和染色体分离的D组，及用于信号转导机制的T组，它们都是LUCA最小蛋白质组的一部分（附录15.1），都基于远古六菌或远古八菌，但它们都不被包括在Ⅰ～Ⅵ的任何最小模型中。因此LUCA并不是前基因组合或最小细胞。作为前生物阶段的最后一个生命体，亦是第一个现代生命体，LUCA显然已经进化超过了最小细胞的阶段。利用其他方法进行的分析也显示，LUCA比最小的生命更为复杂[57,58]。

15.6.2 DNA的出现

有些病毒用DNA基因组，有些是用RNA基因组，也就是说DNA和RNA都是合适的遗传材料。不过所有的细胞基因组都是DNA。虽然DNA作为遗传材料比RNA具有明显优点（见9.3节），但DNA的采用需要发展出复杂的机制。最小的424个COG的LUCA蛋白质组包含24个L组COG用于DNA复制、重组和修复，26 K组COG用于转录，及B，D，F和R组等其他DNA相关的COG。这样，多于10%的LUCA最小蛋白质组都贡献给了DNA信息机制。这需要一个持久而强有力的进化推动力，才能发展出这么多新的COG，足以使RNA基因转化为DNA基因。Forterre提出，这种转化可能是由于RNA生命体需要应对RNA病毒的攻击而转化成DNA生命体[59,60]，但这种病毒防御推动力并不是没有问题的：

(1) 细胞需要经历许多代才能发展出多个COG的DNA机制，很可能没有足够的时间来完成DNA机制来应付尾随而来的RNA病毒。

(2) 就算在许多代之后成功转化为DNA生命体，又怎样能够阻止RNA病毒自身转化为DNA并继续攻击呢？我们知道HIV这样的RNA病毒灵活多变，是最危险的敌人。

综上所述，我们也就有必要考虑除了病毒防御之外的其他动力。一个重要的动力可能是对超高温度条件的适应。正如图15.9所示，因为RNA在超高温度下的半衰期太短，RNA基因组会迅速失去生存能力。因此，生命体要由较冷区域侵入深海热液口[61]，能否成功适应可能就取决于DNA基因的发展。随着生命体越来越接近热液口，这种适应是一个渐进的过程，因此可能经历很多代才完成，从而有足够时间发展出所需的多个COG DNA机制。引导DNA出现的这种热液适应机制可能对LUCA的产生起了关键作用。

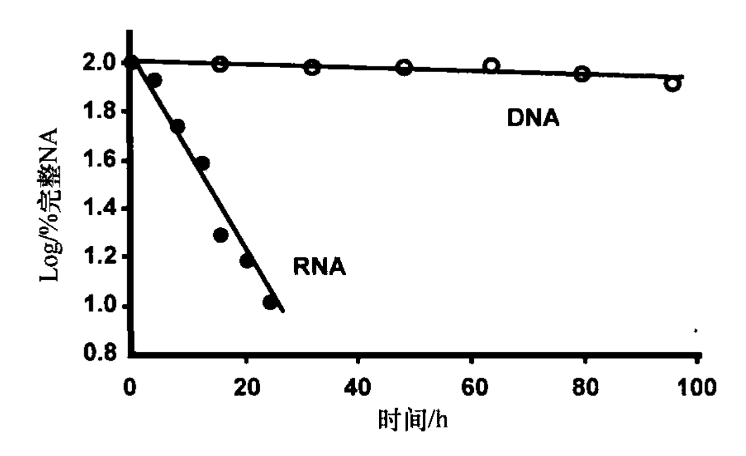


图15.9 RNA和DNA在100℃，pH 7的热降解。所用的RNA序列为gauucaaucugaucucgaugaag，相应的DNA序列为gattcaatctgatctcgatgaag。

15.7 火热之旅

15.7.1 热栖息地

根据最佳生长温度，生命体可被分为4个耐热等级[62]：嗜冷生物4℃、嗜温生物39℃、嗜热生物60℃，以及超嗜热生物90℃或更高。第一个细胞也会在这4个热范围之一形成。

自从在深海热液口发现了生物群落，这些热液口就被认为是生命起源的可能地点[63]。在超高温起源（hyperthermophilic origin，HYO）景象中，这些热液口释放的地热能可以为前生物进化提供适宜的环境。热液口处含有大量的二氧化碳和氢气，在热液温度下的碳固定放能[64]及一系列有机物的前生物合成（见1.3节），可为HYO自养起源提供良好条件。但是，在前生物进化伊始，高温有严

重的缺点，即必要分子的热稳定性[65,66]、在没有复制酶稳定下较短 RNA 或 RNA 类似寡聚物碱基配对的热熔，以及缺乏催化剂为不稳定中间体提供热保护的情况。

为了减少高温带来的困难，并保留热液口生物合成的优势，有人提出生命可能不是在热液口本身处形成，而是在其附近约 90℃的地方形成[67,68]。这种高温起源（thermophilic origin，THO）景象可以更好地平衡自养的可能性及热起源的危险。可是模板 DNA 与一段长 20 个碱基的 DNA 的结合，其熔解温度仅为 50℃。因此，无论在 HYO 还是 THO 条件下，都难以在模板上将单体加入长度小于 20 碱基的 RNA 或 RNA 类似物。因此不难理解，遗传密码结构，大气氨基酸合成及陨石氨基酸这三个与前生物氨基酸可用性相关的证据，都共同证明生命的异养起源，而不是自养起源（见 1.4 节）。

在这种情况下，第一个活细胞可能是中温起源（mesophilic origin，MEO）或嗜冷起源（psychrophilic origin，PSO）。PSO 更利于嘌呤合成及 RNA 多聚化，但在 PSO 条件下所有化学反应的速率都要低于 MEO 条件。不过这两种情况基本都符合异养起源。Thomas Malthus 在《人口论》中早就指出了异养起源的困境，对前卢加（pre-Luca）及后卢加（post-Luca）生命同样适用，即包括人类在内的所有物种都难以无休止地持续高速率繁殖，因为迟早都会遇上缺乏足够的食物的问题，"必然，这种绝对必要的普适自然法则将它们限制在规定的范围内……在动物和植物中，它的效应是种子的耗损、疾病和夭折"[69,70]。前卢加生命和现今的生命一样，倾向于分布到新的生境中去。我们知道，在 36 亿年前 DNA-蛋白质世界初期（图 14.1），最先由遗传密码编码的氨基酸都是来自于环境中的 1 期氨基酸（见 1.4 节），也就是说，在那个时期从环境中仍能得到那些氨基酸。但是，前卢加生命的指数增殖迟早会超过环境中有机物的线性积累。要解决迫在眉睫的碳固定危机，生命体更会加速向新生境的散播。

长期解决碳固定危机的方法是自养。现今，通过碳固定，生命世界完全依赖于自养系统，包括依赖光的光合作用和较少见的依赖氢的产甲烷作用或产乙酸作用。前卢加生命有必要发展出这类的系统。

15.7.2　热液口的生命

生命体做的每一件事都有很强的生物学意义，它们不做的事也同样有意义。古菌擅于自养，可以进行产甲烷作用、产乙酸作用、有像绿硫细菌中的逆柠檬酸循环，以及 Calvin 循环进行碳固定。Calvin 循环是细菌和真核生物中最普遍应用的自养途径，利用 Rubisco 酶通过羟基丙酸盐或羟基酪酸盐途径进行[62,71~73]。虽然适盐古菌也发展了将 ATP 合成与光激活质子泵及细菌视紫红质偶联[74,75]，但是在过去的 30 亿年内，古菌仍然没有发展出成熟的产氧或不产氧光合作用。

细菌利用酯脂，而古菌利用醚脂。在光合作用中，要还原 CO_2，需要 NADH（或 NADPH）在－0.32V 的还原[62]。这种负电位在植物中由光系统产生，在紫细菌中则由反转电子流产生。如果能确定醚脂膜是否有困难产生这个水平的负电位，或古菌有什么其他不利于光合作用的特征，那会是非常有价值的。

无论出于什么原因，古菌并没有发展出基于光合作用的碳固定，而且在白蚁后肠之外的环境中，产甲烷作用比产乙酸作用有竞争力。因此，产甲烷作用应是类似古菌的前卢加生物用以克服面临的碳固定危机的最优出路。这些生命体首先会散播到热液口附近，以吸取那里的非生物合成有机物为生。随着对所需营养源的竞争越来越激烈，它们要以越来越接近热液口来胜出。在这个过程中它们必须提高自己的蛋白质耐热性。而且，随着它们接近热液口，DNA 比 RNA 更好的耐热性成为决定性的选择优势，具有 DNA 基因组的前卢加生物就会占据热液口。较大的小行星体对早期地球的撞击对于原始生物群来说是一个威胁。据估计，直径 440km，大小相当于 Vesta 和 Pallas 的抛射物的撞击，将使海洋沸腾或接近沸腾[76]。这类灾难将进一步促进 DNA 基因组引导的向热液口的迁移。在热液口处，具有 DNA 基因组的前卢加生命发展出产甲烷作用，在足量的 CO_2 和 H_2 中旺盛生长，产生了大量甲烷，令即使在微弱太阳时期的古代地球也避免了冰川时代[77]，完成了 20 种氨基酸的全生命遗传密码，并形成了类似甲烷火菌的超嗜热古菌 LUCA。

一个长期困扰人们的问题是，一个单一的 LUCA 世系，是如何将它的遗传密码赋予所有生命而成为全生命遗传密码，在这个过程中淘汰了所有利用其他遗传密码的世系[78]？任何时候，栖息于不同温度带的生命体间的直接竞争都很有限，那么一种使用最终的通用遗传密码的超嗜热菌，怎样能淘汰所有使用不同遗传密码的嗜冷生物的呢？在热液口发展出 DNA 基因组的 LUCA 给出了一个答案。当 LUCA 的异养后代从热液口散播到较冷的区域时，它们携带了 DNA 基因组及新的通用遗传密码。跨界遇到的异养竞争者带有过时的遗传密码及容易出错的 RNA 基因组。他们之间的竞争是彻底的错配，LUCA 的后代轻易淘汰所有异养竞争者。正如罗马帝国横跨短短几个世纪的武力征服就确立了拉丁语在大范围的使用，现今生命对 4 字母的 DNA 及 20 字母的氨基酸语言的使用也要归功于热液口出现的 LUCA 的强大生化力量。

现代生命的三阶段跨热带起源（Hot-Cross Origin，HCO）设想：①前卢加生命由较冷区域进入热液口；②在水深火热的热液口产生超嗜热产甲烷的 LUCA；③LUCA 的异养后代再次返回较冷区域，确立了它们在地球所有生境中的无他地位，强调了 LUCA 起源与热适应性结合的重要。HCO 描述了第一个嗜冷-嗜温细胞怎样走向热液口的超嗜热产甲烷 LUCA；为 RNA 基因组向 DNA 基因组的转换提供了强大的进化动力；并且解释了全生命密码及其他标准分子生物

装置如何在DNA-蛋白质世界中统一通用至今。候鸟交叉往返于大陆寻找食物、鲑鱼由海洋洄游至山溪进行繁殖、甲烷嗜热菌遍布于各处深海热液口、LUCA的祖先和后代则循相反方向跨越不同温度带——生命体为求生存，什么样的距离都不算远，什么样的旅途都不算艰巨。

15.8 结论

生命起源初始于膜、代谢作用和复制子的整合，发展到LUCA出现的生物界一统的高峰。寻找LUCA是生物学中一个几乎不能够实现的目标，这有两个原因。第一，地球在历史上曾经历了由缺氧大气向富氧大气的转变，带来了众多的环境变化，如冰川期的出现和消失，因此，地球上几乎所有生态环境都经历了极大的改变。栖息在这些地方的有机物别无选择，只能逐渐适应，并在这个过程中失去了许多原始特征，可是，若在一些基因组中没有足够的原始特征延续下来，那就不可能了解LUCA的性质。幸好深海热液口是地球上一直保存着的生境，其气体组成、温度和压力在生物历史上都非常稳定。这使得极端保守的物种如甲烷嗜热菌能够进化得异常缓慢，从而保留了原始特性，对揭示LUCA的性质做出很大的贡献。第二，由于在基因组中缺乏DNA或核糖体RNA旁系同源物，这些分子不能用于寻找LUCA。同时，蛋白质序列又进化得太快，并易受伪迹影响，难以信赖，这样就只剩下tRNA有可能具有合适的信息量来寻找LUCA。令人惊异的是，tRNA的进化是以缓慢而稳定的速率进行，一些tRNA旁系同源物在30亿年间分离得很少。这种进化速率像是为寻找LUCA而量身定做的，而极端保守基因组的tRNA为这种寻找提供了最可靠的引导。这与缺失基因、原始蛋白质、代谢作用及地球化学的证据一起，多重确认了LUCA最接近甲烷嗜热菌的位置。

细菌域和真核生物域由古菌的进化很可能从一开始就带来实质上的更新。比如，细菌和真核生物策略性地利用多种反密码子组合去读取密码子盒，很可能令翻译机制更好地适应变化的环境。同样，真核生物的核膜使它们能够具有多个染色体和较大的基因组。所有已知古菌中，Mac具有最大的基因组，为5.75 Mb，它也是已知唯一可能进入多细胞阶段的古菌[79,80]，这种相关意味着要达到多细胞化，可能至少需要约5.75 Mb的基因组。真核生物中核膜的产生使得远大于5.75 Mb的基因组得以出现，为多细胞生命的出现铺平了道路。将LUCA定位于古菌域，为探索LBACA和LECA的起源提供了基点。在这方面，tRNA序列也是有价值的指针。将所有真核生物tRNA序列与所有细菌tRNA序列进行单独比较，在98万对中只有8个是超高似的、遗传距离小于0.10，而且所有这8个都是蓝藻细菌与Ath的匹配。这个结果不仅证实植物中一些起源于蓝藻细菌的tRNA曾由叶绿体向细胞核迁移，也证实利用高似和超高似tRNA对来寻找

不同生物域间的原始联系是精确的。利用这种方法探测古菌-细菌及真核-古菌关系，指出 LBACA 最接近起源于属泉古菌的气火菌附近的海栖热袍菌，而LECA则源于铁原体类的古菌与立克次氏体或其他细菌通过胞内共生作用形成。

马尔萨斯的种群动力学严格限制了没有类似植物支持动物的受自养支持的异养生命，因此，虽然早期微生物是异养主导，自养生物迟早会出现。由于类似古菌的前卢加生命没有光合作用，产甲烷作用的产生及产甲烷 LUCA 在热液口的出现就十分自然了。并且，正如跨热带起源说景象所描述的那样，由最早的嗜冷-嗜温生命到超嗜热 LUCA 的过程，以及之后 LUCA 的后代返回地球上所有热生态位，一定经历了横跨温度带的过程。因此热适应性这一生化性质就成为 LUCA 阶段原始生命的基础。

太空生物学探索通过太阳系内外的行星和卫星来获得信息，地球上的前生物进化在这方面将是一个很有价值的模型。由于地球生命对水的依赖及对产甲烷 LUCA 的确认，水和甲烷就成为在宇宙中任何位置寻找生命的关键分子。毫无疑问，前生物进化和太空生物学将在未来更好地互相补充完善。今后实在有太多探索在等待明天的前生物学家和太空生物学家来完成。

附录 15.1 LUCA 基因组含有的直系同源聚类（COG）。古八菌中含有的 COG 用加粗字体表示；古六菌中含有而古八菌中缺乏的用普通字体；斜体表示最少蛋白质组 I～VI 中不含有；下划线表示 Mka，Mth 与 Mja 共有的产甲烷途径 COG

类别	序号	直系同源聚类
(1) Translation, ribosomal structure and biogenesis 翻译，核糖体结构和生物起源	**8**	Glutamyl-tRNA synthetase 谷氨酰-tRNA 合成酶
	9	Putative translation factor (SUA5) 假定的翻译因子 (SUA5)
	12	Predicted GTPase, probable translation factor 预计的 GTPase 可能是翻译因子
	13	Alanyl-tRNA synthetase 丙氨酰-tRNA 合成酶
	16	Phenylalanyl-tRNA synthetase alpha subunit 苯丙氨酰-tRNA 合成酶 α 亚基
	17	Aspartyl-tRNA synthetase 天冬氨酰-tRNA 合成酶
	18	Arginyl-tRNA synthetase 精氨酰-tRNA 合成酶
	23	Translation initiation factor 1 (eIF-1/SUI1) and related proteins 翻译起始因子 1 (eIF-1/SUI1) 及相关蛋白
	24	Methionine aminopeptidase 蛋氨酸氨基肽酶
	30	Dimethyladenosine transferase (rRNA methylation) 二甲基腺苷转移酶 (rRNA 甲基化作用)
	48	Ribosomal protein S12 核糖体蛋白 S12
	49	Ribosomal protein S7 核糖体蛋白 S7
	51	Ribosomal protein S10 核糖体蛋白 S10

续表

类别	序号	直系同源聚类
	52	Ribosomal protein S2 核糖体蛋白 S2
	60	Isoleucyl-tRNA synthetase 异亮氨酰-tRNA 合成酶
	72	Phenylalanyl-tRNA synthetase beta subunit 苯丙氨酰-tRNA 合成酶 β 亚基
	80	Ribosomal protein L11 核糖体蛋白 L11
	81	Ribosomal protein L1 核糖体蛋白 L1
	87	Ribosomal protein L3 核糖体蛋白 L3
	88	Ribosomal protein L4 核糖体蛋白 L4
	89	Ribosomal protein L23 核糖体蛋白 L23
	90	Ribosomal protein L2 核糖体蛋白 L2
	91	Ribosomal protein L22 核糖体蛋白 L22
	92	Ribosomal protein S3 核糖体蛋白 S3
	93	Ribosomal protein L14 核糖体蛋白 L14
	94	Ribosomal protein L5 核糖体蛋白 L5
	96	Ribosomal protein S8 核糖体蛋白 S8
	97	Ribosomal protein L6P/L9E 核糖体蛋白 L6P/L9E
	98	Ribosomal protein S5 核糖体蛋白 S5
	99	Ribosomal protein S13 核糖体蛋白 S13
	100	Ribosomal protein S11 以上皆类似核糖体蛋白 S11
	101	Pseudouridylate synthase 假尿苷酸合成酶
	102	Ribosomal protein L13 核糖体蛋白 L13
	103	Ribosomal protein S9 核糖体蛋白 S9
	124	Histidyl-tRNA synthetase 组氨酰-tRNA 合成酶
	130	Pseudouridine synthase 假尿苷合成酶
	143	Methionyl-tRNA synthetase 甲硫氨酰-tRNA 合成酶
	162	Tyrosyl-RNA synthetase 酪氨酰-tRNA 合成酶
	172	Seryl-tRNA synthetase 丝氨酰-tRNA 合成酶
	180	Tryptophanyl-tRNA synthetase 色氨酰-tRNA 合成酶
	182	Predicted translation initiation factor 2B subunit，eIF-2B α/β/δ family 预测的翻译起始因子 2B 亚基，eIF-2B α/β/δ 家族
	184	Ribosomal protein S15P/S13E 核糖体蛋白 S15P/S13E
	185	Ribosomal protein S19 核糖体蛋白 S19
	186	Ribosomal protein S17 核糖体蛋白 S17
	197	Ribosomal protein L16/L10E 核糖体蛋白 L16/L10E

续表

类别	序号	直系同源聚类
	198	Ribosomal protein L24 核糖体蛋白 L24
	199	Ribosomal protein S14 核糖体蛋白 S14
	200	Ribosomal protein L15 核糖体蛋白 L15
	231	Translation elongation factor P（EF-P）/translation initiation factor 5A 翻译延长因子 P（EF-P）/翻译起始因子 5A
	244	Ribosomal protein L10 核糖体蛋白 L10
	252	L-asparaginase/archaeal Glu-tRNAGln amidotransferase subunit D 天冬酰胺酶/古细菌 Glu-tRNAGln 氨基转移酶 D 亚基
	255	Ribosomal protein L29 核糖体蛋白 L29
	256	Ribosomal protein L18 核糖体蛋白 L18
	343	Queuine/archaeosine tRNA-ribosyltransferase 喹啉/古嘌苷 tRNA 核糖基转移酶
	361	Translation initiation factor 1（IF-1）翻译起始因子 1（IF-1）
	423	Glycyl-tRNA synthetase（class II）甘氨酰-tRNA 合成酶（II 类）
	441	Threonyl-tRNA synthetase 苏氨酰基-tRNA 合成酶
	442	Prolyl-tRNA synthetase 脯氨酰基-tRNA 合成酶
	480	Translation elongation factors（GTPases 翻译延长因子（GTPase））
	495	Leucyl-tRNA synthetase 亮氨酰-tRNA 合成酶
	522	Ribosomal protein S4 and related proteins 核糖体蛋白 S4 及相关蛋白
	525	Valyl-tRNA synthetase 缬氨酰基-tRNA 合成酶
	532	Translation initiation factor *2*（IF-*2*；GTPase）翻译起始因子 *2*（IF-*2*；GTPase）
	621	2-Methylthioadenine synthetase 2-甲基硫腺嘌呤合成酶
	1093	Translation initiation factor 2，alpha subunit（eIF-2alpha）翻译起始因子 2，α 亚基（eIF-2α）
	1258	Predicted pseudouridylate synthase 预测的假尿苷酸合成酶
	1325	Predicted exosome subunit 预测的外来体亚基
	1358	Ribosomal protein HS6-type（S12/L30/L7a）核糖体蛋白 HS6 型（S12/L30/L7a）
	1369	RNase P/RNase MRP subunit POP5 RNase P/RNase MRP 亚基 POP5
	1383	Ribosomal protein S17E 核糖体蛋白 S17E
	1384	Lysyl-tRNA synthetase（class 1）赖氨酰基-tRNA 合成酶（I 类）
	1471	Ribosomal protein S4E 核糖体蛋白 S4E
	1491	Predicted RNA-binding protein 预测的 RNA 结合蛋白

续表

类别	序号	直系同源聚类
	1498	Protein implicated in ribosomal biogenesis Nop56p homolog 与核糖体生物合成的 Nop56p 同系物相关的蛋白
	1499	NMD protein affecting ribosome stability and mRNA decay NMD 蛋白影响核糖体稳定性和 mRNA 的降解
	1500	Predicted exosome subunit 预测的外来体亚基
	1503	Peptide chain release factor 1（eRF1）多肽链释放 1 因子（eRF1）
	1514	2′-5′ RNA ligase 2′-5′ RNA 连接酶
	1534	Predicted RNA-binding protein containing KH domain possibly ribosomal protein 预测的 RNA 结合蛋白，包含 KH 结构域，可能包含核糖体蛋白
	1549	Queuine tRNA-ribosyltransferases，contain PUA domain 喹啉 tRNA 核糖基转移酶，包含 PUA 结构域
	1552	Ribosomal protein L40E 核糖体蛋白 L40E
	1588	RNase P/RNase MRP subunit p29 RNase P/RNase MRP p29 亚基
	1601	Translation initiation factor 2，beta subunit（eIF-2beta）/eIF-5 N-terminal domain 翻译起始因子 2，β 亚基（eIF-2β）eIF-5 N 端结构域
	1603	RNase P/RNase MRP subunit p30 RNase P/RNase MRP p30 亚基
	1631	Ribosomal protein L44E 核糖体蛋白 L44E
	1632	Ribosomal protein L15E 核糖体蛋白 L15E
	1676	tRNA splicing endonuclease tRNA 剪接内切酶
	1717	Ribosomal protein L32E 核糖体蛋白 L32E
	1727	Ribosomal protein L18E 核糖体蛋白 L18E
	1736	Diphthamide synthase subunit DPH2 白喉酰胺合成酶 DPH2 亚基
	1746	tRNA nucleotidyltransferase（CCA-adding enzyme）tRNA 核苷酸转移酶（CCA 添加酶）
	1798	Diphthamide biosynthesis methyltransferase 白喉酰胺生物合成甲基转移酶
	1841	Ribosomal protein L30/L7E 核糖体蛋白 L30/L7E
	1867	Dimethylguanosine tRNA methyltransferans 二甲基鸟苷 tRNA 甲基转移酶
	1889	Fibrillarin-like rRNA methylase 类核仁纤维蛋白 rRNA 甲基酶
	1890	Ribosomal protein S3AE 核糖体蛋白 S3AE
	1911	Ribosomal protein L30E 核糖体蛋白 L30E
	1976	Translation initiation factor 6（eIF-6）翻译起始因子 6（eIF-6）
	1997	Ribosomal protein L37AE/L43A 核糖体蛋白 L37AE/L43A
	1998	Ribosomal protein S27AE 核糖体蛋白 S27AE

续表

类别	序号	直系同源聚类
	2004	Ribosomal protein S24E 核糖体蛋白 S24E
	2007	Ribosomal protein S8E 核糖体蛋白 S8E
	2016	Predicted RNA-binding protein（contains PUA domain）预测的 RNA 结合蛋白（包含 PUA 结构域）
	2023	RNase P subunit RPR2 RNase P RPR2 亚基
	2051	Ribosomal protein S27E 核糖体蛋白 S27E
	2053	Ribosomal protein S28E/S33 核糖体蛋白 S28E/S33
	2058	Ribosomal protein L12E/L44/L45/RPP1/RPP2 核糖体蛋白 L12E/L44/L45/RPP1/RPP2
	2075	Ribosomal protein L24E 核糖体蛋白 L24E
	2092	Translation elongation factor EF-1beta 翻译延长因子 EF-1β
	2097	Ribosomal protein L31E 核糖体蛋白 L31E
	2125	Ribosomal protein S6E（S10）核糖体蛋白 S6E（S10）
	2126	Ribosomal protein L37E 核糖体蛋白 L37E
	2139	Ribosomal protein L21E 核糖体蛋白 L21E
	2147	Ribosomal protein L19E 核糖体蛋白 L19E
	2157	Ribosomal protein L20A（L18A）核糖体蛋白 L20A（L18A）
	2163	Ribosomal protein L14E/L6E/L27E 核糖体蛋白 L14E/L6E/L27E
	2167	Ribosomal protein L39E 核糖体蛋白 L39E
	2174	Ribosomal protein L34E 核糖体蛋白 L34E
	2238	Ribosomal protein S19E（S16A）核糖体蛋白 S19E（S16A）
	2260	Predicted Zn-ribbon RNA-binding protein 预测的锌带 RNA 结合蛋白
	2263	Predicted RNA methylase 预测的 RNA 甲基酶
	2511	Archaeal Glu-tRNAGln amidotransferase subunit E（contains GAD domain）古细菌 Glu-tRNAGln 氨基转移酶 E 亚基（含 GAD 结构域）
	2519	tRNA（1-methyladenosine）methyltransferase and related methyltransferases tRNA（1-甲基腺苷）甲基转移酶及相关甲基转移酶
	2888	Predicted Zn-ribbon RNA-binding protein with a function in translation 预测的具有翻译功能的锌带 RNA 结合蛋白
	2890	Methylase of polypeptide chain release factors 多肽链释放因子甲基酶
	3277	RNA-binding protein involved in rRNA processing 参与 rRNA 加工的 RNA 结合蛋白
	5256	Translation elongation factor EF-1alpha（GTPase）翻译延长因子 EF-1α（GTPase）
	5257	Translation initiation factor 2，gamma subunit（eIF-2gamma；GTPase）翻译起始因子 2，γ 亚基（eIF-2γ；GTPase）

续表

类别	序号	直系同源聚类
(A) RNA processing and modification RNA 加工及修饰	*430*	RNA 3′-terminal phosphate cyclase RNA 3′端磷酸环化酶
	2136	Predicted exosome subunit/U3 small nucleolar ribonucleoprotein (snoRNP) component，contains IMP4 domain 预测的外来体亚基/U3 小核仁核蛋白（snoRNP）化合物，含 IMP4 结构域
(K) Transcription 转录	**85**	DNA-directed RNA polymerase，beta subunit/140 kD subunit DNA 介导的 RNA 聚合酶 β 亚基/140kD 亚基
	86	DNA-directed RNA polymerase，beta subunit/160 kD subunit DNA 介导的 RNA 聚合酶 β 亚基/160kD 亚基
	195	Transcription elongation factor 转录延长因子
	202	DNA-directed RNA polymerase，alpha subunit/40 kD subunit DNA 介导的 RNA 聚合酶 α 亚基/40kD 亚基
	250	Transcription antiterminator 转录抗终止因子
	640	Predicted transcriptional regulators 预测的转录调节器
	864	Predicted transcriptional regulators containing the CopG/Arc/MetJ DNA-binding domain and a metal-binding domain 预测的转录调节器，包含 CopG/Arc/Met J DNA 结合结构域及一个金属结合域
	1095	DNA-directed RNA polymerase，subunit E′ DNA 介导的 RNA 聚合酶 E′ 亚基
	1243	Histone acetyltransferase 组蛋白乙酰转移酶
	1293	Predicted RNA-binding protein homologous to eukaryotic snRNP 预测的 RNA 结合蛋白，与真核 snRNP 为同源物
	1308	Transcription factor homologous to NACalpha-BTF3 与 NACα-BTF3 同源的转录因子
	1321	Mn-dependent transcriptional regulator 锰依赖转录调节器
	1378	Predicted transcriptional regulators 预测的转录调节器
	1395	Predicted transcriptional regulator 预测的转录调节器
	1405	Transcription initiation factor TFIIIB，Brf1 subunit 转录起始因子 TFIIIB，Brf1 亚基
	1522	Transcriptional regulators 转录调节器
	1581	Archaeal DNA-binding protein 古细菌 DNA 结合蛋白
	1644	DNA-directed RNA polymerase，subunit N (RpoN/RPB10) DNA 介导的 RNA 聚合酶 N 亚基（RpoN/RPB10）
	1675	Transcription initiation factor IIE，alpha subunit 转录起始因子 IIE，α 亚基
	1758	DNA-directed RNA polymerase，subunit K/omega DNA 介导的 RNA 聚合酶，K/ω 亚基

续表

类别	序号	直系同源聚类
	1761	DNA-directed RNA polymerase，subunit L DNA 介导的 RNA 聚合酶，L 亚基
	1813	Predicted transcription factor，homolog of eukaryotic MBF1 预测的转录因子，与真核 MBF1 同源
	1846	Transcriptional regulators 转录调节器
	1996	DNA-directed RNA polymerase，subunit RPC10（contains C4-type Zn-finger）DNA 介导的 RNA 聚合酶，RPC10 亚基（含 C4 型锌指）
	2012	DNA-directed RNA polymerase，subunit H，RpoH/RPB5 DNA 介导的 RNA 聚合酶，H 亚基，RpoH/RPB5
	2093	DNA-directed RNA polymerase，subunit E'' DNA 介导的 RNA 聚合酶，E''亚基
	2101	TATA-box binding protein（TBP），component of TFIID and TFIIIB TATA 盒结合蛋白（TBP），TFIID 及 TFIIB 的化合物
（L）Replication，recombination and repair 复制，重组及修复	84	Mg-dependent DNase 镁依赖 DNase
	164	Ribonuclease HII 核糖核酸酶 HII
	177	Predicted EndoIII-related endonuclease 预测的 EndoIII 相关内切酶
	258	$5'$-$3'$ exonuclease（including N-terminal domain of PolI）$5'$-$3'$核酸外切酶（包括 PolI 结构域 N 端）
	270	Site-specific DNA methylase 位点特异性 DNA 甲基酶
	350	Methylated DNA-protein cysteine methyltransferase 甲基化 DNA-蛋白半胱氨酸甲基转移酶
	358	DNA primase（bacterial type）DNA 引物酶（细菌型）
	417	DNA polymerase elongation subunit（family B）DNA 聚合酶延长亚基（B 家族）
	419	ATPase involved in DNA repair 参与 DNA 修复的 ATPase
	420	DNA repair exonuclease DNA 修复外切酶
	468	RecA/RadA recombinase RecA/RadA 重组酶
	470	ATPase involved in DNA replication 参与 DNA 复制的 ATPase
	550	Topoisomerase IA 拓扑异构酶 IA
	592	DNA polymerase sliding clamp subunit（PCNA homolog）DNA 聚合酶滑行夹亚基（PCNA 同源物）
	608	Single-stranded DNA-specific exonuclease 单链 DNA 特异性核酸外切酶
	1041	Predicted DNA modification methylase 预测的 DNA 修饰甲基酶

续表

类别	序号	直系同源聚类
	1107	Archaea-specific Recl-like exonuclease，contains Dnal-type Zn finger domain 古细菌特异性类 RecJ 核酸外切酶，含 DnaJ 型锌指结构域
	1111	ERCC4-like helicases 类 ERCC4 解旋酶
	1112	Superfamily 1DNA and RNA helicases and helicase subunits 超家族 1DNA 及 RNA 解旋酶及解旋酶亚基
	1241	Predicted ATPase involved in replication control，Cdc46/Mcm family 预测的 ATPase，参与复制控制，Cdc46/Mcm 家族
	1311	Archaeal DNA polymerase II，small subunit/DNA polymerase delta，subunit B 古细菌 DNA 聚合酶 II，小亚基/DNA 聚合酶 δ，B 亚基
	1389	DNA topoisomerase VI，subunit B DNA 拓扑异构酶 VI，B 亚基
	1423	ATP-dependent DNA ligase，homolog of eukaryotic ligase III ATP 依赖 DNA 连接酶，与真核连接酶 III 同源
	1467	Eukaryotic-type DNA primase，catalytic（small）subunit 真核型 DNA 引物酶，催化（小）亚基
	1525	Micrococcal nuclease（thermonuclease）homologs 微球菌核酸酶（热核酸酶）同源物
	1591	Holiday junction resolvase——archaeal type Holliday 交叉解离酶——古细菌型
	1599	Single-stranded DNA-binding replication protein A（RPA），large（70kD）subunit and related ss DNA-binding proteins 单链 DNA 结合复制蛋白 A（RPA），大（70kD）亚基及相关 ssDNA 结合蛋白
	1637	Predicted nuclease of the RecB family 预测的 RecB 家族核酸酶
	1697	DNA topoisomerase VI，subunit A DNA 拓扑异构酶 VI，A 亚基
	1793	ATP-dependent DNA ligase ATP 依赖 DNA 连接酶
	1933	Archaeal DNA polymerase II，large subunit 古细菌 DNA 聚合酶 II，大亚基
	1948	ERCC4-type nuclease ERCC4 型核酸酶
	2219	Eukaryotic-type DNA primase，large subunit 真核型 DNA 引物酶，大亚基
（B）Chromatin structure and dynamics 染色质结构及动力学	***123***	Deacetylases，including yeast histone deacetylase and acetoin utilization protein 去乙酰化酶，包含酵母组蛋白去乙酰化酶和乙偶姻利用蛋白
	2036	Histones H3 and H4 组蛋白 H3 和 H4

续表

类别	序号	直系同源聚类
(D) Cell cycle control, cell division, chromosome partitioning 细胞周期控制，细胞分裂，染色体分离	***37***	Predicted ATPase of the PP-loop superfamily implicated in cell cycle control 预测的 PP 环超家族 ATPase，参与细胞周期控制
	206	Cell division GTPase 细胞分裂 GTPase
	455	ATPases involved in chromosome partitioning 与染色体分离相关的 ATPase
	489	ATPases involved in chromosome partitioning 与染色体分离相关的 ATPase
	1192	ATPases involved in chromosome partitioning 与染色体分离相关的 ATPase
	1718	Serine/threonine protein kinase involved in cell cycle control 丝氨酸/苏氨酸蛋白激酶，参与细胞周期控制
(T) Signal transduction mechanisms 信号转导机制	***467***	RecA-superfamily ATPases implicated in signal transduction RecA 超家族 ATPase，参与信号转导
	589	Universal stress protein UspA and related nucleotide-binding proteins 全生命弹性蛋白 UspA 及相关核苷酸连接蛋白
	3642	Mn^{2+}-dependent serine/threonine protein kinase 锰离子依赖丝氨酸/苏氨酸蛋白激酶
(M) Cell wall/membrane/envelope biogenesis 细胞壁/膜/封套生物合成	***438***	Glycosyltransferase 糖基转移酶
	449	Glucosamine 6-phosphate synthetase 6-磷酸葡糖胺合成酶
	451	Nucleoside-diphosphate-sugar epimerase 二磷酸核苷糖差向异构酶
	463	Glycosyltransferases involved in cell wall biogenesis 与细胞壁生物合成相关的糖基转移酶
	472	UDP-N-acetylmuramyl pentapeptide phosphotransferase/UDP-*N*-acetylglucosamine-1-1phosphate transferase UDP-*N*-乙酰胞壁酰五肽磷酸转移酶/UDP-*N*-乙酰氨基葡萄糖-1-磷酸转移酶
	668	Small-conductance mechanosensitive channel 小转导机械感应通道
	750	Predicted membrane-associated Zn-dependent proteases 预测的膜相关锌依赖蛋白酶
	794	Predicted sugar phosphate isomerase involved in capsule formation 预测的与囊泡形成相关的糖磷酸异构酶
	1208	Nucleoside-diphosphate-sugar pyrophosphorylase 二磷酸核苷糖焦磷酸化酶
(U) Intracellular trafficking, secretion and vesicular transport 胞内运输，分泌及囊泡转运	**201**	Preprotein translocase subunit SecY 蛋白质前体转位酶，Y 亚基
	341	Preprotein translocase subunit SecF 蛋白质前体转位酶，F 亚基
	342	Preprotein translocase subunit SecD 蛋白质前体转位酶，D 亚基
	541	Signal recognition particle GTPase 信号识别粒子 GTPase
	552	Signal recognition particle GTPase 信号识别粒子 GTPase
	681	Signal peptidase I 信号肽酶 I
	1400	Signal recognition particle 19kDa protein 信号识别粒子 19kDa 蛋白

续表

类别	序号	直系同源聚类
	1989	Type II secretory pathway，prepilin signal peptidase Pul O II 型分泌途径，四型信号肽酶 Pul O
	2064	Flp pilus assembly protein TadC Flp 菌毛装配蛋白 TadC
	2443	Preprotein translocase subunit Sss1 蛋白质前体转位酶 Sss1 亚基
	4962	Flp pilus assembly protein，ATPase CpaF Flp 纤毛装配蛋白 ATPase CpaF
(O) Posttranslational modification，protein turnover，chaperones 翻译后修饰，蛋白质反转，伴侣蛋白	*68*	Hydrogenase maturation factor 氢化酶成熟因子
	71	Molecular chaperone (small heat shock protein) 分子伴侣（小热休克蛋白）
	298	Hydrogenase maturation factor 氢化酶成熟因子
	309	Hydrogenase maturation factor 氢化酶成熟因子
	330	Membrane protease subunits，stomatin/prohibitin homologs 细胞膜蛋白酶亚基，溴化丙胺太林/抑制素同源物
	396	ABC-type transport system involved in Fe-S cluster assembly，ATPase component ABC 型转运系统参与铁-硫簇的装配，ATPase 化合物
	409	Hydrogenase maturation factor 氢化酶成熟因子
	459	Chaperonin GroEL (HSP60 family) 伴侣蛋白 GroEL（HSP60 家族）
	464	ATPases of the AAA+class AAA+类的 ATPase
	492	Thioredoxin reductase 硫氧还原蛋白还原酶
	501	Zn-dependent protease with chaperone function 带有伴侣功能的锌依赖蛋白酶
	533	Metal-dependent proteases with possible chaperone activity 可能有伴侣活性的金属依赖蛋白酶
	555	ABC-type sulfate transport system，permease component ABC 型硫酸盐转运系统，透性酶化合物
	602	Organic radical activating enzymes 有机基激活酶
	638	20S proteasome，alpha and beta subunits 20S 蛋白酶体，α 和 β 亚基
	1047	FKBP-type peptidyl-prolyl cis-trans isomerases 2 FKBP 型蛋白与肽基脯氨酰顺反异构酶 2
	1067	Predicted ATP-dependent protease 预测的 ATP 依赖蛋白酶
	1180	Pyruvate-formate lyase-activating enzyme 丙酮酸甲酸裂解酶激活酶
	1222	ATP-dependent 26S proteasome regulatory subunit ATP 依赖 26S 蛋白酶体调节亚基
	1370	Prefoldin，molecular chaperone implicated in de novo protein folding，alpha subunit 前折叠素，分子伴侣，参与蛋白质重新折叠，α 亚基
	1382	Prefoldin，chaperonin cofactor 前折叠素，伴侣蛋白辅助因素
	1730	Predicted prefoldin，molecular chaperone implicated in de novo protein folding 预测的前折叠素，分子伴侣蛋白参与蛋白质重新折叠
	1899	Deoxyhypusine synthase 脱氧辅蛋白合酶
	2518	Protein-L-isoaspartate carboxylmethyltransferase L-蛋白异天冬氨酸羧甲基转移酶

续表

类别	序号	直系同源聚类
(C) Energy production and conversion 能量产生及转化	**39**	Malate/lactate dehydrogenase 苹果酸盐/乳酸盐脱氢酶
	45	Succinyl-CoA synthetase，beta subunit 琥珀酰-CoA 合成酶，β 亚基
	74	Succinyl-CoA synthetase，alpha subunit 琥珀酰-CoA 合成酶，α 亚基
	243	Anaerobic dehydrogenases，typically selenocysteine-containing 厌氧脱氢酶，尤其是含有硒代半胱氨酸的
	247	Fe-S oxidoreductase 铁-硫氧化还原酶
	371	Glycerol dehydrogenase and related enzymes 甘油脱氢酶及相关酶
	473	Isocitrate/isopropylmalate dehydrogenase 异柠檬酸盐/异丙基苹果酸脱氢酶
	479	Succinate dehydrogenase/fumarate reductase，Fe-S protein subunit 琥珀酸盐脱氢酶/延胡索酸盐还原酶，铁-硫蛋白亚基
	543	2-Polyprenylphenol hydroxylase and related flavodoxin oxidoreductases 2-聚异戊二烯苯酚脱氢酶及相关黄素氧还蛋白氧化还原酶
	636	F0F1-type ATP synthase，subunit c/Archaeal/vacuolar-type H^+-ATPase，subunit K F0F1 型 ATP 合成酶，c 亚基/古细菌/液泡型 H^+-ATPase，K 亚基
	644	Dehydrogenases (flavoproteins) 脱氢酶（黄素蛋白）
	650	Formate hydrogenlyase subunit 4 甲酸盐脱氢酶 4 亚基
	674	Pyruvate：ferredoxin oxidoreductase and related 2-oxoacid：ferredoxin oxidoreductases，alpha subunit 丙酮酸盐：铁氧还蛋白氧化还原酶及相关 2-酮酸：铁氧还蛋白氧化还原酶，α 亚基
	680	Ni，Fe-hydrogenase maturation factor 镍-铁脱氢酶成熟因子
	716	Flavodoxins 黄素氧还蛋白
	731	Fe-S oxidoreductases 铁-硫氧化还原酶
	778	Nitroreductase 硝基还原酶
	1012	NAD-dependent aldehyde dehydrogenases NAD 依赖乙醛脱氢酶
	1013	Pyruvate：ferredoxin oxidoreductase and related 2-oxoacid：ferredoxin oxidoreductases，beta subunit 丙酮酸盐：铁氧还蛋白氧化还原酶及相关 2-酮酸：铁氧还蛋白氧化还原酶，β 亚基
	1014	Pyruvate：ferredoxin oxidoreductase and related 2-oxoacid：ferredoxin oxidoreductases，gamma subunit 丙酮酸盐：铁氧还蛋白氧化还原酶及相关 2-酮酸：铁氧还蛋白氧化还原酶，γ 亚基
	1029	Formylmethanofuran dehydrogenase subunit B 甲酰甲烷呋喃脱氢酶 B 亚基
	1032	Fe-S oxidoreductase 铁-硫氧化还原酶
	1035	Coenzyme F420-reducing hydrogenase，beta subunit 辅酶 F420-还原氢化酶，β 亚基

续表

类别	序号	直系同源聚类
	1036	Archaeal flavoproteins 古细菌黄素蛋白
	1053	Succinate dehydrogenase/fumarate reductase，flavoprotein subunit 琥珀酸盐脱氢酶/延胡索酸盐还原酶，黄素蛋白亚基
	1142	Fe-S-cluster-containing hydrogenase components 2 铁-硫簇抑制氢化酶化合物 2
	1144	Pyruvate：ferredoxin oxidoreductase and related 2-oxoacid：ferredoxin oxidoreductases，delta subunit 丙酮酸盐：铁氧化还原蛋白氧化还原酶及相关的 2-酮酸：铁氧化还原蛋白氧化还原酶，δ 亚基
	1145	Ferredoxin 铁氧化还原蛋白
	1146	Ferredoxin 铁氧化还原蛋白
	1148	Heterodisulfide reductase，subunit A and related polyferredoxins 异二硫化物还原酶 A 亚基及相关多聚铁氧还蛋白
	1149	MinD superfamily P-loop ATPase containing an inserted ferredoxin domain MinD 超家族 P-环 ATPase 包含一个嵌入铁氧化还原蛋白结构域
	1150	Heterodisulfide reductase，subunit C 异二硫化物还原酶 C 亚基
	1151	6Fe-6S prismane cluster-containing protein 6 铁 -6 硫棱柱烷簇含蛋白质
	1152	CO dehydrogenase/acetyl-CoA synthase alpha subunit CO 脱氢酶/乙酰基-CoA 合酶 α 亚基
	1153	Formylmethanofuran dehydrogenase subunit D 甲酰甲烷呋喃脱氢酶 D 亚基
	1155	Archaeal/vacuolar-type H^+-ATPase subunit A 古细菌/析稀胶粒型 H^+-ATPase A 亚基
	1156	Archaeal/vacuolar-type H^+-ATPase subunit B 古细菌/析稀胶粒型 H^+-ATPase B 亚基
	1229	Formylmethanofuran dehydrogenase subunit A 甲酰甲烷呋喃脱氢酶 A 亚基
	1249	Pyruvate/2-oxoglutarate dehydrogenase complex，dihydrolipoamide dehydrogenase（E3）component and related enzymes 丙酮酸盐/2-酮戊二酸脱氢酶复合体，二氢硫辛酰胺脱氢酶（E3）化合物以及相关酶
	1269	Archaeal/vacuolar-type H^+-ATPase subunit 1 古细菌/析稀胶粒型 H^+-ATPase 亚基
	1304	L-lactate dehydrogenase（FMN-dependent）and related alpha-hydroxy acid dehydrogenases L-乳酸盐脱氢酶（FMN-依赖）以及相关 α-羟基酸性脱氢酶
	1390	Archaeal/vacuolar-type H^+-ATPase subunit E 古细菌/液泡型 H^+-ATPase E 亚基
	1394	Archaeal/vacuolar-type H^+-ATPase subunit D 古细菌/液泡型 H^+-ATPase D 亚基

续表

类别	序号	直系同源聚类
	1436	Archaeal/vacuolar-type H^+-ATPase subunit F 古细菌/液泡型 H^+-ATPase F 亚基
	1456	CO dehydrogenase/acetyl-CoA synthase gamma subunit (corrinoid Fe-S protein) CO 脱氢酶/乙酰-CoA 合成酶 γ 亚基（类咕啉铁-硫蛋白）
	1527	Archaeal/vacuolar-type H^+-ATPase subunit C 古细菌/液泡型 H^+-ATPase C 亚基
	1592	Rubrerythrin 红素氧还蛋白
	1614	CO dehydrogenase/acetyl-CoA synthase beta subunit CO 脱氢酶/乙酰 CoA 合成酶 β 亚基（类咕啉铁-硫蛋白）
	1625	Fe-S oxidoreductase，related to NifB/MoaA family 铁-硫氧还蛋白，与 NifB/MoaA 家族相关
	1819	Glycosyl transferases，related to UDP-glucuronosyltransferase 糖基转移酶，与 UDP 葡萄糖醛酸转移酶相关
	1880	CO dehydrogenase/acetyl-CoA synthase epsilon subunit CO 脱氢酶/乙酰-CoA 合成酶 ε 亚基（类咕啉铁-硫蛋白）
	1838	Tartrate dehydratase beta subunit/Fumarate hydratase class I，C-terminal domain 酒石酸盐脱氢酶 β 亚基/I 类延胡索酸盐脱氢酶 C 端结构域
	1908	Coenzyme F420-reducing hydrogenase，delta subunit 辅酶 F420 还原脱氢酶，δ 亚基
	1927	Coenzyme F420-dependent N (5)，N (10) -methenyltetrahydromethanopterin dehydrogenase 辅酶 F420 依赖 N (5)，N (10) -甲酰基四氢甲基蝶呤脱氢酶
	1941	Coenzyme F420-reducing hydrogenase，gamma subunit 辅酶 F420 还原脱氢酶，γ 亚基
	1951	Tartrate dehydratase alpha subunit/Fumarate hydratase class I，N-terminal domain 酒石酸盐脱氢酶 α 亚基/I 类延胡索酸盐脱氢酶 N 端结构域
	2037	Formylmethanofuran：tetrahydromethanopterin formyltransferase 甲酰甲烷呋喃：四氢甲基蝶呤甲酰转移酶
	2048	Heterodisulfide reductase，subunit B 异二硫化物还原酶 B 亚基
	2055	Malate/L-lactate dehydrogenases 苹果酸盐/L-乳酸盐脱氢酶
	2069	CO dehydrogenase/acetyl-CoA synthase delta subunit (corrinoid Fe-S protein) CO 脱氢酶/乙酰-CoA 合成酶 δ 亚基（类咕啉铁-硫蛋白）
	2141	Coenzyme F420-dependent N5，N10-methylene tetrahydromethanopterin reductase and related flavin-dependent oxidoreductases 辅酶 F420 依赖 N5，N10 亚甲基四氢甲基蝶呤还原酶及相关黄素依赖氧化还原酶

续表

类别	序号	直系同源聚类
	2218	Formylmethanoturan dehydrogenase subunit C 甲酰甲烷呋喃脱氢酶 C 亚基
	2221	Dissimilatory sulfite reductase (desulfovirdin), alpha and beta subunits 异化的亚硫酸盐还原酶（脱硫绿胺霉素），α 和 β 亚基
	2710	Nitrogenase molybdenum-iron protein, alpha and beta chains 固氮酶钼-铁蛋白质，α 和 β 链
	3259	Coenzyme F420-reducing hydrogenase, alpha subunit 辅酶 F420-还原脱氢酶，α 亚基
	3260	Ni, Fe-hydrogenase III small subunit 镍，铁-脱氧酶 III 小亚基
	3261	Ni, Fe-hydrogenase III large subunit 镍，铁-脱氧酶 III 大亚基
	4074	H2-forming N5, N10-methylenetetrahydromethanopterin dehydrogenase 形成 H2 N5，N10-亚甲基四氢甲基蝶呤脱氢酶
(G) Carbohydrate transport and metabolism 二氧化碳转运及代谢	**57**	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase 甘油醛-3-磷酸盐脱氢酶/4-磷酸赤藓糖脱氢酶
	61	Predicted sugar kinase 预测的糖激酶
	63	H2-forming N5, N10-methylenetetrahydromethanopterin dehydrogenase 形成 H2 N5，N10-亚甲基四氢甲基蝶呤脱氢酶
	120	Ribose 5-phosphate isomerase 核酸糖 5-磷酸盐异构酶
	126	3-Phosphoglycerate kinase 3-磷酸甘油酸盐激酶
	148	Enolase 烯醇酶
	149	Triosephosphate isomerase 磷酸丙糖异构酶
	235	Ribulose-5-phosphate 4-epimerase and related epimerases and aldolases 核酮糖-5-磷酸盐 4-差向异构酶，相关差向异构酶以及醛缩酶
	269	3-Hexulose-6-phosphate synthase and related proteins 3-已酮糖-6-磷酸盐合酶以及相关蛋白质
	483	Archaeal fructose-1, 6-bisphosphatase and related enzymes of inositol monophosphatase family 古细菌 1，6-二磷酸果糖及相关单磷酸肌醇家族的酶
	524	Sugar kinases, ribokinase family 糖激酶，核糖激酶家族
	574	Phosphoenolpyruvate synthase/pyruvate phosphate dikinase 磷酸烯醇丙酮酸盐合酶/丙酮酸盐磷酸盐
	662	Mannose-6-phosphate isomerase 甘露糖-6-磷酸盐异构酶
	1082	Sugar phosphate isomerases/epimerases 糖磷酸盐异构酶/差向异构酶
	1109	Phosphomannomutase 磷酸甘露糖变位酶
	1363	Cellulase M and related proteins 纤维素酶 M 及相关蛋白

续表

类别	序号	直系同源聚类
	1830	DhnA-type fructose-1，6-bisphosphate aldolase and related enzymes DhnA 型 1，6-二磷酸果糖醛缩酶及相关酶
	1980	Archaeal fructose 1，6-bisphosphatase 古细菌 1，6-二磷酸果糖
	2074	2-phosphoglycerate kinase 2-磷酸甘油酸激酶
	3635	Predicted phosphoglycerate mutase，AP superfamily 预测的磷酸甘油酸变位酶，AP 超家族
	3839	ABC-type sugar transport systems，ATPase components ABC 型糖转运系统，ATPase 化合物
(E) Amino acid transport and metabolism 氨基酸转运及代谢	***2***	Acetylglutamate semialdehyde dehydrogenase 乙酰谷氨酸半醛脱氢酶
	6	Xaa-Pro aminopeptidase Xaa-Pro 氨基肽酶
	10	Arginase/agmatinase/formimionoglutamate hydrolase，arginase family 精氨酸酶/凝集酶/亚胺甲基谷氨酸水解酶，精氨酶酶家族
	65	3-lsopropylmalate dehydratase large subunit 3-异丙基苹果酸脱氢酶大亚基
	66	3-lsopropylmalate dehydratase small subunit 3-异丙基苹果酸脱氢酶小亚基
	75	Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase 丝氨酸-丙酮酸盐氨基转移酶/古细菌天冬氨酸氨基转移酶
	76	Glutamate decarboxylase and related PLP-dependent proteins 谷氨酸盐脱羧酶及相关 PLP 依赖蛋白
	78	Ornithine carbamoyltransferase 鸟氨酸甲氨酰转移酶
	79	Histidinol-phosphate/aromatic aminotransferase and cobyric acid decarboxylase 组胺醇-磷酸盐/芳香族氨基转移酶及钴啉胺酸脱羧酶
	111	Phosphoglycerate dehydrogenase and related dehydrogenases 磷酸甘油酸脱氢酶及相关脱氢酶
	112	Glycine/serine hydroxymethyltransferase 甘氨酸羟甲基转移酶
	119	lsopropylmalate/homocitrate/citramalate synthases 异丙基苹果酸/高柠檬酸/柠苹酸合成酶
	136	Aspartate-semialdehyde dehydrogenase 天冬氨酸-半醛脱氢酶
	174	Glutamine synthetase 谷氨酰胺合成酶
	289	Dihydrodipicolinate reductase 二氢甲基吡啶酸还原酶
	329	Dihydrodipicolinate synthase/N-acetylneuraminate lyase 二氢甲基吡啶酸合成酶/乙酰神经氨酸裂解酶
	367	Asparagine synthase (glutamine-hydrolyzing) 氨羰丙氨酸合酶（谷氨酸-水解）
	436	Aspartate/tyrosine/aromatic aminotransferase 天冬氨酸/酪氨酸/芳香族化合物转氨酶

续表

类别	序号	直系同源聚类
	440	Acetolactate synthase，small（regulatory）subunit 乙酰乳酸合酶，小（调节）亚基
	460	Homoserine dehydrogenase 高丝氨酸脱氢酶
	498	Threonine synthase 苏氨酸合酶
	527	Aspartokinases 天冬氨酸激酶
	548	Acetylglutamate kinase 乙酰谷氨酸激酶
	560	Phosphoserine phosphatase 磷酸丝氨酸磷酸酶
	620	Methionine synthase II（cobalamin-independent）蛋氨酸合酶 II（钴维生素-自主）
	1387	Histidinol phosphatase and related hydrolases of the PHP family 组氨醇磷酸酶以及相关 PHP 家族水解酶
	1812	Archaeal S-adenosylmethionine synthetase 古细菌 S 腺苷甲硫氨酸合成酶
	4992	Ornithine/acetylornithine aminotransferase 鸟氨酸/乙酰鸟氨酸转氨酶
（F）Nucleotide transport and metabolism 核苷酸转运与代谢	***5***	Purine nucleoside phosphorylase 嘌呤核苷磷酸化酶
	15	Adenylosuccinate lyase 腺苷酸基琥珀酸裂解酶
	34	Glutamine phosphoribosyl-pyrophosphate amidotransferase 谷氨酸磷酸核糖-焦磷酸盐酰胺转移酶
	41	Phosphoribosyl-carboxy-aminoimidazole mutase 磷酸核糖-羧基-氨基咪唑变位酶
	44	Dihydroorotase and related cyclic amidohydrolases 二氢乳清酸酶以及相关循环氨基水解酶
	46	Phosphoribosylformyl glycinamidine synthase，synthetase domain 磷酸核糖甲酸基甘氨脒合酶，合成酶结构域
	47	Phosphoribosylformylglycinamidine synthase，glutamine amidotransferase domain 磷酸核糖甲酸基甘氨脒合酶，谷氨酸酰胺转移酶结构域
	104	Adenylosuccinate synthase 腺苷酸基琥珀酸合酶
	105	Nucleoside diphosphate kinase 核苷二磷酸盐激酶
	125	Thymidylate kinase 胸苷酸激酶
	127	Xanthosine triphosphate pyrophosphatase 三磷酸黄嘌呤核苷焦磷酸酶
	150	Phosphoribosylaminoimidazole synthetase 磷酸核糖氨基咪唑合成酶
	151	Phosphoribosylamine-glycine ligase 磷酸核糖胺-甘氨酸连接酶
	152	Phosphoribosyl-aminoimidazole-succinocarboxamide synthase 磷酸核糖-氨基咪唑-琥珀羧胺合成酶
	167	Dihydroorotate dehydrogenase 二氢乳清酸脱氢酶

续表

类别	序号	直系同源聚类
	284	Orotidine-5′-phosphate decarboxylase 5′-磷酸乳清苷脱羧酶
	402	Cytosine deaminase and related metal-dependent hydrolases 胞嘧啶脱氨酶及相关金属依赖水解酶
	461	Orotate phosphoribosyltransferase 乳清酸盐磷酸核糖转移酶
	462	Phosphoribosylpyrophosphate synthetase 磷酸核糖焦磷酸盐合成酶
	503	Adenine/guanine phosphoribosyltransferases and related PRPP-binding proteins 腺嘌呤/鸟嘌呤磷酸核糖转移酶及相关 PRPP 结合蛋白
	504	CTP synthase CTP 合成酶
	516	IMP dehydrogenase/GMP reductase IMP 脱氢酶/GMP 还原酶
	518	GMP synthase——Glutamine amidotransferase domain GMP 合成酶——谷氨酰胺氨基转移酶结构域
	519	GMP synthase，PP-ATPase domain/subunit GMP 合成酶，PP-ATPase 结构域/亚基
	528	Uridylate kinase 尿苷酸激酶
	540	Aspartate carbamoyltransferase，catalytic chain 天冬氨酸甲氨酰转移酶催化链
	717	Deoxycytidine deaminase 脱氧胞嘧啶脱氨酶
	1051	ADP-ribose pyrophosphatase 核糖-ADP 焦磷酸酶
	1102	Cytidylate kinase 胞苷酸激酶
	1328	Oxygen-sensitive ribonucleoside-triphosphate reductase 氧敏感性三磷酸核糖核苷还原酶
	1437	Adenylate cyclase，class 2（thermophilic）2 类腺苷酸环化酶（嗜热）
	1618	Predicted nucleotide kinase 预测的核苷酸激酶
	1781	Aspartate carbamoyltransferase，regulatory subunit 天冬氨酸甲氨酰转移酶调节亚基
	1828	Phosphoribosyl-formylglycinamidine synthase，PurS component 磷酸核糖-甲酰甘氨脒合成酶，PurS 化合物
	1936	Predicted nucleotide kinase（related to CMP and AMP kinases）预测的核苷酸激酶（与 CMP 及 AMP 激酶相关）
	2019	Archaeal adenylate kinase 古细菌腺嘌呤核苷酸激酶
（H）Coenzyme transport and metabolism 辅酶转运及代谢	***43***	3-Polyprenyl-4-hydroxybenzoate decarboxylase and related decarboxylases 3-聚异戊二烯-4-羟苯酸盐脱羧酶及相关脱羧酶
	142	Geranylgeranyl pyrophosphate synthase 香叶酰香叶酰焦磷酸盐合成酶
	157	Nicotinate-nucleotide pyrophosphorylase 烟酸盐-核苷酸焦磷酸化酶
	163	3-Polyprenyl-4-hydroxybenzoate decarboxylase 3-聚异戊二烯-4-羟苯酸盐脱羧酶
	171	NAD synthase NAD 合成酶
	214	Pyridoxine biosynthesis enzyme 吡哆醇生物合成酶

续表

类别	序号	直系同源聚类
	237	Dephospho-CoA kinase 脱磷酸辅酶 A 激酶
	294	Dihydropteroate synthase and related enzymes 二氢叶酸合成酶及相关酶
	301	Thiamine biosynthesis ATP pyrophosphatase 硫胺素生物合成 ATP 焦磷酸酶
	303	Molybdopterin biosynthesis enzyme 亚钼嘌呤生物合成酶
	311	Predicted glutamine amidotransferase involved in pyridoxine biosynthesis 预测的谷氨酰胺氨基转移酶参与维生素 B_6 生物合成
	315	Molybdenum cofactor biosynthesis enzyme 钼辅因子生物合成酶
	351	Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase 羟甲基嘧啶/磷甲基嘧啶激酶
	368	Cobalamin-5-phosphate synthase 5-磷酸-钴胺素合成酶
	379	Quinolinate synthase 喹啉酸合成酶
	382	4-Hydroxybenzoate polyprenyltransferase and related prenyltransferases 4-羟苯酸盐聚异戊二烯转移酶及相关异戊二烯转移酶
	452	Phosphopantothenoylcysteine synthetase/decarboxylase 磷酸泛酰半胱氨酸合成酶/脱羧酶
	499	S-adenosylhomocysteine hydrolase 腺苷高半胱氨酸水解酶
	521	Molybdopterin biosynthesis enzymes 亚钼嘌呤生物合成酶
	611	Thiamine monophosphate kinase 硫胺素单磷酸盐激酶
	720	6-Pyruvoyl-tetrahydropterin synthase 6-丙酮酰四氢蝶呤合成酶
	746	Molybdopterin-guanine dinucleotide biosynthesis protein A 亚钼嘌呤-鸟嘌呤二核苷酸生物合成蛋白 A
	1270	Cobalamin biosynthesis protein CobD/CbiB 钴维生素生物合成蛋白 CobD/CbiB
	1339	Transcriptional regulator of a riboflavin/FAD biosynthetic operon 核黄素/FAD 生物合成操纵子的转录调节器
	1635	Flavoprotein involved in thiazole biosynthesis 黄素蛋白参与噻唑生物合成
	1763	Molybdopterin-guanine dinucleotide biosynthesis protein 亚钼嘌呤-鸟嘌呤二核苷酸生物合成蛋白 A
	1767	Triphosphoribosyl-dephospho-CoA synthetase 三磷酸核糖脱磷酸 CoA 激酶
	2038	NaMN：DMB Phosphoribosyltransferase NaMN：DMB 磷酸核糖转移酶
	2266	GTP：adenosylcobinamide-phosphate guanylyltransferase GTP：腺苷基钴啉醇酰胺磷酸鸟苷酰基转移酶
	2896	Molybdenum cofactor biosynthesis enzyme 钼辅因子生物合成酶

续表

类别	序号	直系同源聚类
(I) Lipid transport and metabolism 脂类转运和代谢作用	***20***	Undecaprenyl pyrophosphate synthase 十一碳二烯焦磷酸合成酶
	170	Dolichol kinase 长醇激酶
	183	Acetyl-CoA acetyltransferase 乙酰-CoA 乙酰转移酶
	575	CDP-diglyceride synthetase CDP-甘油二酯合成酶
	615	Cytidylyltransferase 胞苷酰转移酶
	671	Membrane-associated phospholipid phosphatase 膜结合磷脂磷酸酶
	1257	Hydroxymethylglutaryl-CoA reductase 羟甲基戊二酸-CoA 还原酶
	1267	Phosphatidylglycerophosphatase A and related proteins 磷脂酰甘油磷酸酶 A 及相关蛋白
	1577	Mevalonate kinase 甲羟戊酸激酶
	3425	3-Hydroxy-3-methylglutaryl CoA synthase 3-羟基-3-甲烷戊二酰 CoA 合成酶
(P) Inorganic ion transport and metabolism 无机离子转运和代谢	168	Trk-type K^+ transport systems, membrane components Trk 型钾离子转运系统，膜化合物
	306	Phosphate/sulphate permeases 磷酸盐/硫酸盐透性酶
	370	Fe^{2+} transpot system protein B 铁离子转运系统蛋白 B
	477	Permeases of the major facilitator superfamily 易化扩散载体超家族透性酶
	530	Ca^{2+}/Na+antiporter 钙离子/钠离子反向转运体
	569	K^+ transport systems, NAD-binding component 钾离子转运系统，NAD 结合化合物
	619	ABC-type cobalt transport system, permease component CbiQ and related transporters ABC 型钴转运系统，透性酶化合物 CbiQ 及相关转运体
	704	Phosphate uptake regulator 磷酸摄入调节器
	725	ABC-type molybdate transport system periplasmic component ABC 型钼酸盐转运系统，周质化合物
	1122	ABC-type cobalt transport system, ATPase component ABC 型钴转运系统，ATPase 化合物
	1226	3-Hydroxy-3-methylglutaryl CoA synthase 3-羟基-3-甲基戊二酰 CoA 合成酶
	1918	Fe^{2+} transport system protein A 铁离子转运系统蛋白质 A
(Q) Secondary metabolite biosynthesis, transport and catabolism 二级代谢生物合成，转运及分解作用	***179***	2-Keto-4-pentenoate hydratase/2-oxohepta-3-ene-1, 7-dioic acid hydratase (catechol pathway) 2-酮-4-戊烯酸酯水合酶/2-氧代庚基-3-烯-1，7-二元酸水合酶（儿茶酚途径）
	500	SAM-dependent methyltransferases SAM 依赖甲基转移酶

续表

类别	序号	直系同源聚类
(R) General function prediction only 一般功能，仅为预期	73	EMAP domain EMAP 结构域
	312	Predicted Zn-dependent proteases and their inactivated homologs 预测的锌依赖蛋白酶及其非激活同源物
	375	Zn finger protein HypA/HybF (possibly regulating hydrogenase expression) 锌指蛋白 HypA/HybF（可能调节氢化酶表达）
	433	Predicted ATPase 预测的 ATPase
	446	Uncharacterized NAD (FAD)-dependent dehydrogenases 非典型 NAD (FAD) 依赖脱氢酶
	456	Acetyltransferases 乙酰转移酶
	491	Zn-dependent hydrolases, including glyoxylases 锌指依赖水解酶，包括乙二醛酶
	517	FOG: CBS domain FOG：CBS 结构域
	535	Predicted Fe-S oxidoreductases 预测的 Fe-S 氧化还原酶
	603	Predicted PP-loop superfamily ATPase 预测的 PP 环超家族
	622	Predicted phosphoesterase 预测的磷酸酯酶
	663	Carbonic anhydrases/acetyltransferases, isoleucine patch superiamily 碳酸酐酶/乙酰转移酶，异亮氨酸修补超家族
	714	MoxR-like ATPases 类 MoxR ATPase
	1011	Predicted hydrolase (HAD superfamily) 预测的水解酶（HAD 超家族）
	1019	Predicted nucleotidyltransferase 预测的核苷酸转移酶
	1078	HD superfamily phosphohydrolases HD 超家族磷酸水解酶
	1084	Predicted GTPase 预测的 GTPase
	1094	Predicted RNA-binding protein (contains KH domains) 预测的 RNA 结合蛋白（包含 KH 结构域）
	1100	GTPase SAR1 and related small G proteins GTPase SAR1 及小 G 蛋白
	1163	Predicted GTPase 预测的 GTPase
	1201	Lhr-like helicases 类 Lhr 解旋酶
	1204	Superfamily II helicase II 超家族解旋酶
	1205	Distinct helicase family with a unique C-terminal domain including a metal-binding cysteine cluster 独特的解旋酶家族，具有一个独特的 C 端结构域，包含半胱氨酸金属结合簇
	1234	ATPase components of various ABC-type transport systems, contain duplicated ATPase 不同 ABC 型转运系统中的 ATPase 化合物，包含二倍 ATPase

续表

类别	序号	直系同源聚类
	1235	Metal-dependent hydrolases of the beta-lactamase superfamily β-内酰胺酶超家族 I 中的金属依赖水解酶
	1237	Metal-dependent hydrolases of the beta-lactamase superfamily β-内酰胺酶超家族 II 中的金属依赖水解酶
	1245	Predicted ATPase，RNase Linhibitor（RL1）homolog 预测的 ATPase，RNase L 抑制子（RL1）同源物
	1313	Uncharacterized Fe-S protein pflX，homolog of pyruvate formate lyase activating proteins 非典型 Fe-S 蛋白 pflX，为丙酮酸甲酸裂解酶激活蛋白同源物
	1326	Uncharacterized archaeal Zn-finger protein 非典型古细菌锌指蛋白
	1355	Predicted dioxygenase 预测的加双氧酶
	1365	Predicted ATPase（PP-loop superfamily）预测的 ATPase（PP 环超家族）
	1407	Predicted ICC-like phosphoesterases 预测的类 ICC 磷酸酯酶
	1412	Uncharacterized proteins of PilT N-term superfamily PilT N-期超家族非典型蛋白
	1418	Predicted HD superfamily hydrolase 预测的 HD 超家族水解酶
	1439	Predicted nucleic acid-binding protein，consists of a PIN domain and a Zn-ribbon module 预测的核酸结合蛋白，由 PIN 结构域和锌带模块组成
	1458	Predicted DNA-binding protein containing PIN domain 预测的 DNA 结合蛋白，包含 PIN 结构域
	1537	Predicted RNA-binding proteins 预测的 RNA 结合蛋白
	1545	Predicted nucleic-acid-binding protein containing a Zn-ribbon 预测的核酸结合蛋白，包含一个锌带
	1571	Predicted DNA-binding protein containing a Zn-ribbon domain 预测的 DNA 结合蛋白，包含一个锌带结构域
	1608	Predicted archaeal kinase 预测的古细菌激酶
	1646	Predicted phosphate-binding enzymes，TIM-barrel fold 预测的磷酸结合酶，TIM 桶折叠
	1759	ATP-utilizing enzymes of ATP-grasp superfamily（probably carboligases）ATP 捕获超家族中利用 ATP 的酶（可能是醛连接酶）
	1779	C4-type Zn-finger protein C4 型锌指蛋白
	1782	Predicted metal-dependent RNase，consists of a metallo-beta-lactamase domain and an RNA-binding KH domain 预测的金属依赖 RNase，由一个金属 β-内酰胺酶结构域及一个 RNA 结合 KH 结构域组成

续表

类别	序号	直系同源聚类
	1818	Predicted RNA-binding protein，contains THUMP domain 预测的 RNA 结合蛋白，包含 THUMP 结构域
	1829	Predicted metal-dependent RNase，consists of a metallo-beta-lactamase domain and an RNA-binding KH domain 预测的金属依赖 RNase，由一个金属 β-内酰胺酶结构域及一个 RNA 结合 KH 结构域组成
	1831	Predicted metal-dependent hydrolase（urease superfamiy）预测的金属依赖水解酶（尿素酶超家族）
	1855	ATPase（PilT family）ATPase（PilT 家族）
	1907	Predicted archaeal sugar kinases 预测的古细菌糖激酶
	1938	Archaeal enzymes of ATP-grasp superfamily ATP 捕获超家族中的古细菌酶
	1964	Predicted Fe-S oxidoreductases 预测的 Fe-S 氧化还原酶
	1988	Predicted membrane-bound metal-dependent hydrolases 预测的膜连接金属依赖水解酶
	2047	Uncharacterized protein（ATP-grasp superfamily）非典型蛋白质（ATP 捕获超家族）
	2102	Predicted ATPases of PP-loop superfamily 预测的 PP 环 ATPase 超家族
	2118	DNA-binding protein DNA 结合蛋白
	2129	Predicted phosphoesterases，related to the lcc protein 预测的磷酸酯酶，与 lcc 蛋白相关
	2151	Predicted metal-sulfur cluster biosynthetic enzyme 预测的金属-硫簇生物合成酶
	2220	Predicted Zn-dependent hydrolases of the beta-lactamase fold 预测的 β 内酰胺酶折叠锌依赖水解酶
	2244	Membrane protein involved in the export of O-antigen and teichoic acid 与 O 抗原和磷壁酸输出相关的膜蛋白
	2520	Predicted methyltransferase 预测的甲基转移酶
	3269	Predicted RNA-binding protein，contains TRAM domain 预测的 RNA 结合蛋白，包含 TRAM 结构域
（S）Unknown proteins 未知蛋白		*11*，***62***，*32*，*327*，*392*，***432***，*585*，*1303*，***1371***，*1379*，*1373*，*1422*，*1430*，*1432*，*1433*，***1460***，*1469*，*1578*，*1617*，*1627*，***1628***，***1679***，***1690***，*1698*，***1701***，***1720***，*1784*，***1786***，***1814***，*1833*，*1844*，***1849***，***1888***，***1909***，*1916*，***1931***，*1945*，***1990***，*1991*，***1992***，*2029*，*2034*，***2078***，*2083*，*2450*，*2454*，*2457*，*2892*，

参考文献

扩展阅读

1. Gillis M, Vandamme P, De Vos P et al. Polyphasic taxonomy. In: Boone DR, Castenholz RW, eds. Bergey's Manual of Systematic Bacteriology, 2nd ed. New York:Springer-Verlag, 2001: 43-48.
2. Krane DE, Raymer ML. Fundamental concepts of bioinformatics. Benjamin Cummings, 2003.
3. Salemi M, Vandamme AM, eds. The phylogenetic handbook. Cambrige:Cambridge University Press, 2003.
4. Staley JT, Reysenbach, eds. Biodiversity of microbial life. Wiley-Liss 2002.

专题文献

5. Forterre P, Philippe H. Where is the root of the universal tree of life? BioEssays 1999; 21:871-879.
6. Brinkmann H, Philippe H. Archaea sister group of bacteria? Indications from tree reconstruction artifacts in ancient phylogenies. Mol Biol Evol 1999; 16:817-825.
7. Brown JR, Doolittle WF. Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. Proc Natl Acad Sci USA 1995; 92:2441-2445.
8. Wong JT, Chen J, Mat WK et al. Polyphasic evidence delineating the root of life and roots of biological domains. Gene 2007; 403:39-52.
9. Pennisi E. Is it time to uproot the tree of life? Science 1999; 284:1305-1307.
10. Wolf YI, Rogozin IB, Grishin NV et al. Genome trees and the tree of life. Trends Genet 2002; 18:472-479.
11. Xue H, Tong KL, Marck C et al. Transfer RNA paralogs: evidence for genetic code-amino acid biosynthesis coevolution and an archaeal root of life. Gene 2003; 310:59-66.
12. Xue H, Ng SK, Tong KL et al. Congruence of evidence for a methanopyrus-proximal root of life based on transfer RNA and aminoacyl-tRNA synthetase genes. Gene 2005; 360:120-130.
13. Sun FJ, Caetano-Anolles G. Evolutionary patterns in the sequence and structure of transfer RNA: early origins of archaea and viruses. PLoS Computaional Biol 2008; 4:e1000018.
14. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria and Eucarya. Proc Natl Acad Sci USA 1990; 87:4576-4579.
15. Woese CR. Interpreting the universal phylogenetic tree. Proc Natl Acad Sci USA 2000; 97:8392-8396.
16. Olsen GJ, Woese CR, Overbeek R. The winds of (evolutionary) change: breathing new life into microbiology. J Bacteriol 1994; 176:1-6.
17. Danchin A, Fang G, Noria S. The extant core bacterial proteome is an archive of the origin of life. Proteomics 2007; 7:875-889.
18. Holland M. The unseen extreme: microbiology at deep-sea hydrothermal vents. J Marine Edu 2005; 21:45-48.
19. Eigen M, Winkler-Oswatitsch R. Transfer-RNA, an early gene? Naturwissen 1981; 68:282-292.
20. Florentz C, Giege R. tRNA-like structures in plant viral RNAs. In: Soll D, RajBhandary UL. eds. tRNA Structure, Biosynthesis and Function: Washington DC:ASM Press, 1995:141-163.
21. Maizels N, Weiner AM. The genomic-tag hypothesis: what molecular fossils tell us about the evolution of tRNA. In: Gesteland RF, Cech TR, Atkins JF, eds. The RNA World. Woodbury:Cold Spring Harbor Laboratory Press, 1999:79-111.
22. Tong KL, Wong JT. Anticodon and wobble evolution. Gene 2004; 333:169-177.
23. Simonson AB, Servin JA, Skophammer RG et al. Decoding the genome tree of life. Proc Natl Acad Sci USA 2005; 102:6608-6613.
24. Qin Y, Polacek N, Vesper O et al. The highly conserved LepA is a ribosomal elongation factor that back-translocates the ribosome. Cell 2006; 127:721-733.
25. Bocchetta M, Gribaldo S, Sanangelantoni A et al. Phylogenetic depth of the bacterial Genera Aquifex and Thermotoga inferred from analysis of ribosomal protein, elongation factor and RNA polymerase subunit sequences. J Mol Evol 2000; 50:366-380.
26. Brown JR, Douady CJ, Italia MJ et al. Universal trees based on large combined protein sequence data sets. Nature Genet 2001; 28:281-285.
27. Altschul SF et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nulc Acid Res 1997; 25:3389-3402.
28. Battistuzzi FU, Feijao A, Hedges SB. 2004. A genomic timescale of prokaryote evolution: insight into the origin of methanogenesis, phototrophy and the colonization of land. BMC Evol Biol 2004; 4:44-57.
29. Raymond J, Segre D. The effect of oxygen on biochemical networks and the evolution of complex life. Science 2006; 311:1764-1767.
30. Archetti M, Di Giulio M. The evolution of the genetic code took place in an anaerobic environment. J Theoret Biol 2007; 245:169-174.
31. Di Giulio M. The universal ancestor and the ancestor of bacteria were hyperthermophiles. J Mol Evol 2003; 57:721-730.
32. Di Giulio M. A comparison of proteins of Pyrococcus furiosus and Pyrococcus abyssi: barophily in the physicochemical properties of amino acids and in the genetic code. Gene 2005; 346:1-6.
33. Di Giulio M. Structuring of the genetic code took place at acidic pH. J Theoret Biol 2005; 237:219-226.
34. Stetter KO. Hyperthermophilic prokaryotes. FEMS Microbiol Rev 1996; 18:149-158.
35. Gaucher EA, Govindarajan S, Ganesh OK. Palaeotemperature trend for Precambrian life inferred from resurrected proteins. Nature 2008; 451:704-708.
36. Gouy M, Chaussidon M. Evolutionary biology: ancient bacteria liked it hot. Nature 2008; 451:635-636.
37. Leigh JA. Chapter 4. Evolution of energy metabolism. In: Biodiversity of Microbial Life, eds. Staley JT and Reysenbach A-L. Wiley-Liss 2002; 103-120.
38. Falkowski PG. Tracing oxygen's imprint on Earth's metabolic evolution. Science 2006; 311:1724-1725.
39. Slesarev AI et al. The complete genome of hyperthermophilic Methanopyrus kandleri AV19 and monophyly of archaeal methanogens. Proc Natl Acad Sci USA 2002; 99:4644-4649.
40. Dopson M, Baker-Austin C, Hind A et al. Characterization of Ferroplasma isolates and Ferroplasma acidarmanus sp., extreme acidophiles from acid mine drainage and industrial bioleaching environments, Appl Environ Microbiol 2004; 70:2079-2088.
41. Golyshina OV, Timmis KN. Ferroplasma and relatives, recently discovered cell wall-lacking archaea making a living in exteremely acid, heavy metal-rich environments. Environ Microbiol 2005; 7:1277-1288.
42. Searcy DG, Hixon WG. Cytoskeletal origins in sulfur-metabolizing archaebacteria, BioSystems 1991; 25:1-11.
43. Margulis L, Chapman M, Guerrero R et al. The last eukaryotic common ancestor (LECA): acquisition of cytoskeletal motility from aerotolerant spirochaetes in the Proterozoic Eon. Proc Natl Acad Sci USA 2006; 103:13080-13085.
44. Pisani D, Cotton JA, McInerney JO. Supertrees disentangle the chimeric origin of eukaryotic genomes, Mol Biol Evol 2007; 24:1752-1760.
45. Kurland CG, Collins LJ, Penny D. Genomics and the irreducible nature of eukaryotic cells. Science 2006; 312:1011-1014.
46. Fitzpatrick DA, Creevy CJ, McInerney JO. Genome phylogenies indicate a meaningful proteobacterial phylogeny and support a grouping of the mitochondria with the Rikettsiales. Mol Biol Evol 2006; 23:74-85.
47. Poole AM, Penny D. Evaluating hypotheses for the origin of eukaryotes. BioEssays 2006; 29:74-84.
48. Mat WK, Xue H, Wong JT. The genomics of LUCA. Frontiers in Biosc 2008; 13:5605-5613.
49. Tatusov RL, Koonin EV, Lipman DJ. A genomic perspective on protein families. Science 1997; 278:631-637.
50. Ouzounis CA, Kunin V, Darzentas N et al. A minimum estimate for the gene content of the last universal common ancestor—exobiology from a terrestrial perspective. Res Micobiol 2006; 157:57-68.
51. Woese CR. Evolutionary questions: the "Progenote". Science 1990; 247:789.
52. Nakabachi A, Yamashita A, Toh H et al. The 160-kilobase genome of the bacterial endosymbiont Carsonella. Science 2006; 314:267.
53. Islas S, Becerra A, Luisi PL et al. Comparative genomics and the gene complement of a minimal cell. Orig Life Evol Biosph 2004; 34:243-56.
54. Glass JI, Assad-Garcia N, Alperovich N et al. Essential genes of a minimal bacterium. Proc Natl Acad Sci USA 2006; 103:425-430.
55. Kobayashi K. et al. Essential Bacillus subtilis genes. Proc Natl Acad Sci USA 2003; 100:4678-4683.
56. Danchin A, Fang G, Noria S. The extant core bacterial proteome is an archive of the origin of life. Proteomics 2007; 7:875-889.
57. Ranea JA, Sillero A, Thornton JM et al. Protein superfamily evolution and the Last Universal Common Ancestor (LUCA). J Mol Evol 2006; 63:513-525.
58. Wang M, Yafremava LS, Caetano-Anolles D et al. Reductive evolution of architectural repertoires in proteomes and the birth of the tripartite world. Genome Res 2007; 17:1572-1585.
59. Forterre P. The two ages of the RNA world and the transition to the DNA world: a story of viruses and cells. Biochimie 2005; 87:793-803.
60. Forterre P. Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domains. Proc Natl Acad Sci USA 2006; 103:3669-3674.
61. Islas S, Velasco AM, Becerra A et al. Hyperthermophily and the origin and earliest evolution of life. Int Microbiol 2003; 6:87-94.

62. Madigan MT, Martinko JM, Parker J. Brock Biology of Microorganisms. 10th ed. Prentice Hall 2002; 152:448-559.
63. Corliss JB, Baross JA, Hoffman SE. A hypothesis concerning the relationship between submarine hot springs and the origin of life on Earth. Oceanol Acta 1981; 4:59-69.
64. Shock EL. Chemical environment in submarine hydrothermal systems. Orig Life Evol Biosphere (suppl) 1992; 22:67-107.
65. Miller SL, Bada JL. Submarine hot springs and the origin of life. Nature 1988; 334:609-611.
66. Bada JL, Lazcano A. Origin of life. Some like it hot, but not the first biomolecules. Science 2002; 296:1982-1983.
67. Cody GD, Boctor NZ, Filley TR et al. Primordial carbonylated iron-sulfur compounds and the synthesis of pyruvate. Science 2000; 289:1337-1340.
68. Koonin EV, Martin W. On the origin of genomes and cells within inorganic compartments. Trends in Genet 2005; 21:647-654.
69. Malthus TR. An Essay on the Principle of Population. New York:Norton, 1976 rpt.
70. Larson EJ. On the Origins of Darwinsism Evolution. Modern Library Chronicles, 2004:68.
71. Henstra A, Dijkema C, Stams AJM. Archaeoglobus fulgidus couples CO oxidation to sulfate reduction and acetogenesis with transient formate accumulation. Environ Microbiol 2007; 9:1836-1841.
72. Berg IA, Kockelkorn D, Buckel W et al. A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. Science 2007; 318:1782-1786.
73. Thauer RK. A fifth pathway of carbon fixation. Science 2007; 318:1732-1733.
74. Danon A, Stoeckenius W. Photophosphorylation in Halobacterium halobium. Proc Natl Acad Sci USA 1974; 71:1234-1238.
75. Schafer G, Engelhard M, Muller V. Bioenergetics of the Archaea. Microbiol Mol Biol Rev 1999; 63:570-620.
76. Sleep NH, Zahnle KJ, Kasting JF et al. Annihilation of ecosystems by large asteroid impacts on the early Earth. Nature 1989; 342:139-142.
77. Kasting JF, Siefert JL. Life and the evolution of Earth's atmosphere. Science 2002; 296:1066-1068.
78. Wong JT. The evolution of a universal genetic code. Proc Natl Acad Sci USA 1976; 73:2336-2340.
79. Macario AJL, Conway de Macario E. The molecular chaperon system and other anti-stress mechanisms in archaea. Frontiers Biosc 2001; 6:d262-283.
80. Galagan JE et al. The genome of M. acetivorans reveals extensive metabolic and physiological diversity. Genome Res 2002; 12:532-542.

[General Information]
书名=生命的来历：前生物进化与太空生物学
作者=王子晖等编著
页数=240
SS号=12724775
出版日期=2011